DOCUMENT RESUME

ED 036 670 AA 000 508

TITLE Research Reporting Sections: National Council of Teachers of Mathematics Golden Jubilee Year, 48th Annual Meeting. SMAC Report.

INSTITUTION Ohio State Univ., Columbus. ERIC Information Analysis Center for Science Education.

PUB DATE Mar 70

NOTE 69p.

EDRS PRICE EDRS Price MF-$0.50 HC-$3.55

DESCRIPTORS *Abstracts, College Mathematics, *Conference Reports, Elementary School Mathematics, Instruction, Learning, *Mathematics Education, *Research, Secondary School Mathematics

IDENTIFIERS National Council of Teachers of Mathematics

ABSTRACT

Abstracts of research papers presented at the Research Reporting Sections of the 48th Annual Meeting of the National Council of Teachers of Mathematics were compiled and edited by the ERIC Information Analysis Center for Science and Mathematics Education and are included in this publication. These abstracts describe research involving a wide variety of topics related to mathematics education at the elementary, secondary, and college levels. Topics investigated in these papers include: the relationship of rationalization of conservation and mathematical achievement; the effect of socio-economic differences on the teaching of readiness for mathematical concepts; children's concept of limits; perception of plane sections of solid figures; student evaluation of teachers; the relationship between personality and differential achievement; diagnosis of teacher behavior characteristics; characteristics of students who are successful in problem solving; the effect of different presentations of word problems; the effect of teaching Euclidean geometry via transformations; the use of a developmental model to build a probability unit; performance on certain discovery tasks; understanding certain concepts of logic; comparison of different methods of teaching proof using logic; the feasibility of proof in elementary school; and the role of counterexamples in the development of mathematical concepts. (FL)

RESEARCH REPORTING SECTIONS

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

GOLDEN JUBILEE YEAR

48th ANNUAL MEETING

SMAC/SCIENCE & MATHEMATICS EDUCATION INFORMATION ANALYSIS CENTER

. . . an information center to organize and disseminate information and materials on science and mathematics education to teachers, administrators, supervisors, researchers, and the public. A joint project of The Ohio State University and the Educational Resources Information Center of USOE.

RESEARCH REPORTING SECTIONS

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

GOLDEN JUBILEE YEAR

48th ANNUAL MEETING

Shoreham Hotel
Washington, D.C.
April 1-4, 1970

PREFACE

The ERIC Information Analysis Center for Science and Mathematics Education has compiled abstracts of research papers to be presented at this conference. Some editing was done by the ERIC staff to provide a general format for the abstracts. Special recognition should be given to Dr. F. Joe Crosswhite, Mrs. Maxine Weingarth, Mrs. Cheryl Brosey, Mrs. Cassandra Balthaser, and Miss Susan Hedger who were responsible for compiling and preparing the report.

Many of the papers will be published in journals or be made available through the ERIC system. These will be announced in Research in Education and other publications of the ERIC system.

March, 1970

Robert W. Howe
Director
ERIC Information Analysis
Center for Science and
Mathematics Education
The Ohio State University
1460 West Lane Avenue
Columbus, Ohio 43221

Sponsored by the Educational Resources Information Center of the United States Office of Education and The Ohio State University.

This publication was prepared pursuant to a contract with the Office of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

TABLE OF CONTENTS

PREFACE

Page

PAPERS PRESENTED AT THE MEETING

Research Reporting Section 1

Leader: F. Joe Crosswhite, The Ohio State University, Columbus, Ohio.

Speakers: 1. W. George Cathcart, University of Alberta, Edmonton, Alberta, "The Relationship Between Primary Students' Rationalization of Conservation and Their Mathematical Achievement."

2. E. Harold Harper, University of Colorado, Boulder, Colorado. "The Identification of Socio-Economic Differences and Their Effect on the Teaching of Readiness for 'New Math Concepts' in the Kindergarten."

3. Stanley F. Taback, New York University, New York, New York, "The Child's Concept of Limit."

4. Edward J. Davis, University of Georgia, Athens, Georgia. "A Study of the Ability of Selected School Pupils to Perceive the Plane Sections of Selected Solid Figures."

THE RELATIONSHIP BETWEEN PRIMARY STUDENTS' RATIONALIZATION OF CONSERVATION AND THEIR MATHEMATICAL ACHIEVEMENT

W. George Cathcart
University of Alberta
Edmonton, Alberta

The purpose of the research was three-fold: (1) to determine the frequency with which different kinds of rationalizations were given to justify conservation responses to Piagetian conservation test items, (2) to investigate differences in social and personal characteristics of subjects who preferred different kinds of rationalizations and (3) to examine the relationship between the different modes of rationalization for conservation and achievement in mathematics.

It has been well established that the ability to conserve is correlated in a positive direction to achievement in mathematics. Such a relationship has not been established between different modes of rationalizing conservation and mathematics achievement. If such a relationship exists then instructional strategies should take into account the style of thought exemplified by the most successful mode of rationalization.

The modes of rationalization considered in the present study were those suggested by Jean Piaget. Piaget claims that a subject, when asked why a property remains invariant after an invariant but perceptually distorting transformation, may respond with one of three arguments. A child may respond with a "reversibility" argument by saying, "If you moved it back it would be the same." Piaget calls the second type of rationalization "identity." For example: "You didn't add anything or take anything away so it is the same." In the present study, this type of response was classified as an "operational identity" response to differentiate it from a "substantive identity" rationalization such as "Its the same water (paper, etc.)." The third type of argument which children might use to justify conservation is called "Compensation" by Piaget. A response such as, "This one is longer here but it is shorter here, so it is the same," is indicative of this mode of rationalization for conservation.

A random sample of 120, grade 2 and 3 subjects from 12 schools (generally middle class), was given an eight-item conservation test. In addition, tests were administered to obtain a measure of each child's vocabulary, intelligence, listening ability and mathematics achievement. The conservation and vocabulary tests were administered individually and the remaining tests were administered as group tests in each school.

Some of the major null hypotheses tested were:

1. There is no significant difference in the observed frequency with which each mode of rationalization is chosen and a chance rectangular frequency distribution. This hypothesis was tested with a Kolmogorov-Smirnov one-sample test.

2. There is no significant relationship between the types of conservers (total or partial) and the mode of rationalization expressed for conservation. A chi-square test of independence was used to test this hypothesis.

3. There is no significant difference between subjects exhibiting different modes of rationalization for conservation on the following criteria: (a) intelligence, (b) socio-economic status, (c) vocabulary, (d) listening ability, and (e) conservation. This hypothesis was tested with a one-way analysis of variance with a Newman-Keuls comparison of ordered means used as an <u>a posteriori</u> test of differences.

4. There is no significant relationship between mode of rationalization for conservation and: (a) age, (b) grade, and (c) sex. A chi-square test of independence was used to test this hypothesis.

5. There are no significant main effects due to mode of rationalization on the mathematics achievement test. A two-way analysis of variance was used to test this hypothesis.

6. There are no significant differences among subjects who use only one mode, two different modes, three different modes, or all four modes of rationalization for conservation in mathematical achievement. A one-way analysis of variance with a Newman-Keuls comparison of ordered means was used to test this hypothesis.

It was found that the primary students used in the present study preferred to rationalize conservation with identity arguments. Nearly 80 percent of the subjects were placed in the two identity categories. Total conservers and partial conservers seemed to differ somewhat in the mode of rationalization they used for conservation. Partial conservers tended to use compensation and substantive identity more than total conservers who in turn used operational identity and reversibility to a greater extent than partial conservers. However, both groups used identity-type arguments most extensively.

There were no significant differences among the four rationalization groups in intelligence, socio-economic status, vocabulary, listening ability, age, grade, or sex. The only characteristic in which the groups differed significantly was the ability to conserve. Students who preferred to rationalize conservation with reversibility arguments had the highest score on the conservation test. Students in the compensation group had the lowest mean on the conservation test.

The reversibility group had the highest mean score on the mathematics test, but none of the differences among the groups could be considered significant. However, when the subjects were classified by the number of different modes of rationalization they used, significant differences were observed. Generally, students who used three different modes of rationalization over the whole conservation test were superior in achievement. On the other hand, students who could use only one mode of rationalization were the lowest achievers in mathematics.

The heavy reliance by primary students on identity arguments suggests that this style of thinking should be capitalized on in teaching strategy. For example, identity implies that all the basic facts for a given number should be taught together, i.e., $6 + 1 = 7$; $5 + 2 = 7$; etc. Furthermore, the general trend for the highest achievement scores to be obtained by subjects who rationalized conservation with reversibility arguments suggests that addition and subtraction should be taught together as should multiplication and division since one is the 'reverse' of the other.

Another major implication arises out of the finding that students who can rationalize in several ways are the best achievers in mathematics. This suggests that our approach to teaching should be an approach which emphasizes multiple strategies in problem solving. In other words, reasoning about relationships, problems, etc. should be stressed at the expense of rules.

THE IDENTIFICATION OF SOCIO-ECONOMIC DIFFERENCES AND THEIR EFFECT ON THE TEACHING OF READINESS FOR 'NEW MATH CONCEPTS' IN THE KINDERGARTEN

E. Harold Harper
University of Colorado
Boulder, Colorado

The concepts espoused by Jean Piaget have received acclaim by some and criticism by others. His study of conservation as it relates to mathematical thinking has been the subject of many research studies. The present study is another in the general area of children's abilities to recognize numerical properties of sets of objects.

This project is a continuation of a study conducted under the sponsorship of the Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison, Wisconsin during the Spring of 1967. That study is reported in Technical Report No. 38. It succeeded in teaching conservation of numerousness to small groups of kindergarten children, in a middle-class community, with one highly trained teacher conducting the lessons.

The purpose of the present study was to see if the typical classroom teacher, in schools differing in socio-economic levels and with whom the ultimate value of the treatment rests, can successfully use the lessons developed in the previous study to effect conservation of numerousness with kindergarten children. An additional purpose was to ascertain whether special inservice training of the teachers is necessary for the treatment to have its full educational impact on the pupils.

The 1967 study in Oconomowoc, Wisconsin public schools was successful, but this community does not have many children from low socio-economic families. For this reason, the investigator felt it necessary to use the lessons in a community where adequate samples of all socio-economic levels could be used.

The study of 1967 controlled the teacher variable by having only one teacher (who received instruction and demonstrations for teaching each lesson) conduct the experimental lessons. To make this study more generalizable, teachers were chosen at random and some given training on how to teach the lessons while others proceeded on their own. Also, instructional matarials were used in the 1967 experiment which were consistent for all classes. In the present study the teachers were allowed to use those materials he could obtain most easily. The same materials used in the previous study were specified in the Teachers' Guide to the lessons but the investigator suggested that they might use alternative materials as long as the general procedure and format of the lessons was not violated.

By allowing for these variations, the applicability of the lessons to "normal" classroom situations allowed a better analysis of the desirability of using the lessons with all kindergarten children under the

typical teacher's direction. The following questions were considered.

1. Can the typical classroom teacher teach the conservation lessons as successfully as a specially trained expert? Is special training necessary or beneficial for the treatment to have maximum benefit?

2. Is the treatment of greater value for pupils from disadvantaged backgrounds? (Is there a treatment by socio-economic-status interaction?)

3. Is the treatment of greater value for younger kindergarten children than for older ones? If environmental deprivation has a crystalizing effect, as is often claimed, one might expect to find this to be the case. (Is there a treatment-by-age interaction?)

4. Do younger children from disadvantaged backgrounds, who may have more cognitive flexibility, benefit more from the lessons? (Is there a treatment-by-socio-economic-status-by-age interaction?)

The "Test of Conservation of Numerousness" used in the 1967 study was designed to be used with small groups. A Hoyt-Reliability coefficient of .91 was computed with an item analysis for the instrument. It was also correlated with an individually administered test and a correlation between total scores of .84 was obtained. This test functioned well in the previous study but the investigator felt that the item analysis indicated a need for revising some of the instructions and physical arrangement of figures on a few of the items. For this reason, a revised form was developed and was correlated with the original instrument.

A Special Study Committee of the Denver Public Schools has completed an investigation of equality of educational opportunity in the district and has amassed much data on socio-economic levels in various sections of the school district. Using the results of this study, a random selection of eighteen classrooms was made in areas 1-6 as defined in this study. The three classrooms per area are described below:

1. Control:

   One control classroom which was tested at the end of the experiment. This classroom received no other treatment.

2. Experimental 1:

   One experimental classroom where the teacher received special training by the principal investigator in methods of teaching each of the twelve lessons.

3. Experimental 2:

One experimental classroom where the teacher used his own interpretation of "Lessons for Teaching Conservation of Numerousness to Kindergarten Children" in order to teach the twelve lessons.

The series of twelve weekly thirty-minute lessons developed for the Oconomowoc experiment was duplicated and supplied to all the experimental teachers. The control teachers were not informed that their classes were participating in an experiment.

The experiment was conducted during the spring semester, 1969. The test was administered to the children in both experimental and control classrooms using the revised Form III of the "Test of Conservation of Numerousness." This testing commenced one week after the close of the experiment.

A total of 484 kindergarten students in the City of Denver, Colorado were the subjects in this study. They ranged in age from 65 months to 89 months with a mean age of 73.47 months. The subjects were members of eighteen classrooms picked at random from six geographic areas of the city. The choice of A.M. or P.M. classes was also randomly assigned. The teachers in these classrooms ranged in age from 22 years to 61 years with a mean age of 35.09. Their years of teaching experience ranged from one year to 31 years with a mean of 9.739 years of experience. Thirteen teachers held the B.A. degree, two the B.A. plus X hours, and three had the M.A.

The experimental design was a 3X7X2 factorial design with one covariate. The factors were Treatments (t), Socio-Economic Status of head of family (SES), and Age (A). The dependent measure was the students' scores on the "Test of Conservation of Numerousness," Form III. The covariate was the teacher's years of teaching experience. All of the factors were considered to be fixed factors.

In a preliminary analysis, the teachers' years of teaching experience had some effect and it was used as a covariate in the final analysis of the data. The data were then analyzed by the analysis of covariance on the CDC 6400 with program BMD 05V - Biomedical Computer Program 05V.

Campbell and Stanley's design 6, "The Posttest only Control Group Design" $\begin{pmatrix} R & X & O_1 \\ R & & O_2 \end{pmatrix}$ was employed in this study.

The testing instrument, "Test of Conservation of Numerousness," Form III was correlated with Form 1 of the same test. A Pearson Product Moment parallel forms reliability coefficient of .70 was obtained. The investigator felt that this reliability was substantial and that this form of the test would be easier to administer to the large population involved in this study than would the orginial Form 1.

ANALYSIS OF VARIANCE RESULTS TABLE***

Effect	Degrees of Freedom	F-Value
Treatment (T)	2	3.8619*
SES (S)	6	3.04193**
Age (A)	1	.21105
TXS	12	1.0979
SXA	6	2.6893
TXA	2	2.2047
TXSXA	12	.46653
Mean Square Error	440	15.1306

* $P < .05$
** $P < .01$
*** Only mean square error, degrees of freedom, and F values for effects are given. All other values can be reconstructed from these.

The overall results of the analysis of variance indicate differences in the treatment at the .05 level of confidence. The adjusted mean for Control was 7.0496, for Experimental 1 was 7.4045, and for Experimental 2 was 6.3831. With 2 and 440 degrees of freedom this yields an F value of 3.8619. Employing the Tukey Test, the post hoc comparison of means indicated that the contributing effect was due to the difference between Experimental 1 and Experimental 2. The comparison indicates that Experimental 1 was more effective than Experimental 2 at the .05 level of confidence.

Socio-Economic-Status and SES by Age were also significant at the .01 and .05 levels respectively. Post hoc comparisons indicate that children coming from homes where the "head of family" was employed in the three lowest categories, performed better in the Experimental 1 teaching-learning situation than in either of the other two treatments. The interaction between Socio-Economic-Status and Age was disregarded as it did not involve an interaction with the treatment.

There were four questions this study was designed to answer. They are listed on page 2. The only question which would be answered affirmatively would be question number two,"Is the treatment of greater value for pupils from disadvantaged backgrounds?" This is true if by combining categories 5, 6, and 7 one considers these groups to be lower Socio-Economic groups of disadvantaged backgrounds. Then the Experimental 1 treatment, where the teachers met weekly with the investigator for inservice instruction on the use of the lessons, proved most successful of the three treatments for enhancing conservation of numerousness with children from categories 5, 6, and 7. This approached significance at the .10 level of confidence when the Tukey Test was employed.

The interaction between Treatment and Age approaches significance with an F value or 2.2047 with 6 and 440 degrees of freedom. On closer observation, however, there is no consistent pattern among means.

THE CHILD'S CONCEPT OF LIMIT

Stanley F. Taback
New York University
New York, New York

The concept of limit is fundamental to the study of mathematical analysis. In recent years, mathematics educators have begun to introduce topics in analysis much earlier in the school curriculum. For example, Robert B. Davis of the Madison Project has done considerable work with eighth and ninth grade students on the subject of convergence. The Secondary School Mathematics Curriculum Improvement Study (SSMCIS) has introduced the completeness property of the real number system in terms of the concept of "least upper bond" to eighth grade students. SSMCIS also includes formal definitions of the limit of a sequence, the Cauchy criterion, and the derivative of a function for grades nine and ten. Furthermore, there are strong indications that mathematicians are looking toward the elementary school as a starting point for such instruction. The report of the Cambridge Conference on School Mathematics, Goals for School Mathematics,[1] recommends that the child in grades 3-6 begin to understand the distinction between rational and irrational numbers and to consider infinite sequences of real numbers. As mathematics educators concern themselves with the child's readiness for such instruction, they must, necessarily, consider his awareness of the concept of limit.

Except for this study there exists practically no research related to the child's concept of limit. It is hoped that this investigation of selected concepts basic to the development of limit will contribute significantly to the work of school mathematics curriculum innovators. They are the ones who will be devising materials to introduce the real numbers and related topics in analysis into a mathematically sound elementary or junior high school program.

The purposes of the study are:

1. To seek information concerning the child's intuitive understandings of selected concepts that are included in every situation involving a limiting process. These concepts are:

   I) Functional rule of correspondence

   II) Neighborhood of a point

   III) Convergence (Divergence)

   IV) Limit point

[1]Cambridge Conference on School Mathematics, Goals for School Mathematics (Boston: Houghton Mifflin Company, 1963)

2. To record the development of each of the above concepts across three age levels: 8, 10, and 12 years.

Eight tasks were devised for use in the study. Seven tasks were overtly non-mathematical, though each embodied a limiting process. For example, one task described a rabbit who successively hops halfway towards a given point; another task pictured a sequence of nested magazine covers converging to a point. In these seven tasks, the subject was questioned on the functional rule of correspondence, the convergence or divergence involved, and the properties of the limit point, if one were present. In the remaining task, he was examined solely on the concept of neighborhood.

The population studied was defined as New York City independent-school children. Five schools participated: Columbia Grammar School; Downtown Community School; Elisabeth Irwin School and its elementary division, Little Red School House; New Lincoln School; and Walden School. Five children were selected at random from each school at each of the three age levels in question -- eight years, ten years, and twelve years.

The tasks were administered individually to each subject by the investigator. A tape recording was made of each interview and, subsequently, was evaluated in terms of a predetermined rating scheme. The following five categories were used to rate subject responses: Clear understanding, Some understanding, Uncertain understanding, No understanding, and Evidence lacking. In order to measure the reliability of the rating scheme instrument, the investigator trained a doctoral candidate at Teachers College, Columbia University to rate the child's responses. Based upon ten tape recordings, selected at random and rated separately by each person, there was 93.2 percent of agreement.

The principal findings may be summarized as follows:

Functional rule of correspondence. Two principles were true, in general. First, a maximum of 60 percent of the eight-year-olds could understand any particular rule of correspondence. Second, at least 80 percent of each of the ten and twelve-year-old age groups could do so.

Neighborhood. Given the defining properties of an open interval and an open circle, less than 20 percent of the eight-year-olds could locate the boundary of either neighborhood. Not one of these children could envision more than fifty points within the open interval or a hundred points within the open circle. Sixty percent of each of the ten and twelve-year-olds could locate the boundary of the open interval; 40 percent of each group could do so for the open circle. Only one ten-year-old could envision infinitely many points within a neighborhood; by contrast, eight twelve-year-olds could do so.

Convergence. There was an increase in performance with age on every task. The difference, however, was much greater between the eight- and ten-year-olds than between the ten- and twelve-year-olds. Approximately 70 to 80 percent of the ten- and twelve-year-olds who understood the concept of convergence on the concrete level also understood it on the abstract level. This figure dropped to 40 percent for the eight-year-olds. Those questions that most demanded a level of thought liberated from physical limitations were answered only be twelve-year-olds.

Limit Point. All three age groups performed better on those tasks in which the limit point was not actually visible in the accompanying diagram. However, even in these tasks, less than 20 percent of the eight-year-olds could conceptualize the limit point. At least 50 percent of the twelve-year-olds could do so while the ten-year-olds performed at a slightly lower level. When the limit point was visible, only one eight-year-old and less than 20 percent of each of the two older age groups could conceptualize it. The most difficult limit points were understood only by twelve-year-old boys.

In general, the eight-year-olds could do little more than follow a simple rule of correspondence. The ten- and twelve-year-olds were much more successful on all four concepts examined. However, with few exceptions, only twelve-year-old subjects exhibited the ability to go beyond the perceptible and to operate in terms of theoretical possibilities; consequently, only these subjects were able to conceptualize an infinite process.

A STUDY OF THE ABILITY OF SELECTED SCHOOL PUPILS TO PERCEIVE THE PLANE SECTIONS OF SELECTED SOLID FIGURES

Edward J. Davis
University of Georgia
Athens, Georgia

This study was a modification of Boe's dissertation[1] investigating the responses of school pupils to some of Piaget's sectioning tasks reported in The Child's Conception of Space.[2] The sectioning test consisted of sixteen multiple-choice items devised by Boe. Four cuts were visualized on each of four solids (cube, cone, cylinder and rectangular solid). The task modifications of this study included: 1) a manipulative work period preceding the test, 2) a slight increase in sample and cell size, and 3) a two-level decrease in grade level.

The work period was designed to provide each subject with experience in cross-sectioning of irregular shaped objects, and to familiarize each subject with the test format. Thus, in particular, the work period was intended to increase test validity by putting the subjects at ease and by providing familiarity with the type of situation they were going to be asked to visualize. The work period consisted of individual sectioning activities and numerous visitations by the researcher with each student. In each visitation a dialogue was established concerning the tasks at hand.

Mastery of sectioning is a critical task in Piaget's theory of the development of representational space in children. Piaget maintains the child's mastery of sectioning solid figures indicates both the emerging dominance of Euclidean concepts and the necessary correspondence between Euclidean and projective relationships in representational thought. Cross-sections appear in the middle- and senior-high school mathematics curriculum as aids in determining volume of three-dimensional geometric solids and also as models of some second-degree equations studied in secondary school algebra and geometry courses. If, as Boe reported, children in grades eight, ten, and twelve have not mastered cross-sectioning, the use of cross-sections would appear to confuse students rather than add meaning to the mathematical situation being studied.

---

1 Barbara L. Boe, "A Study of the Ability of Secondary School Pupils to Perceive the Plane Sections of Selected Solid Figures," Mathematics Teacher, Vol. LXI, April, 1968.

2 Jean Piaget and Barbel Inhelder, The Child's Conception of Space, W. W. Norton and Company Inc., New York, 1967, pp. 251-267.

Since the method of gathering data included both a Piagetian dialogue type situation and the structured testing situation of Boe, it was felt that the modifications could yield results which might help explain some of the differences in Piaget and Boe's findings in children's ability to perceive sections. Piaget reported eleven-year-old's as having mastered conic sections. Boe reported that sectioning ability is incomplete through grade twelve and that sectioning ability is related to mental ability instead of to age.

Ninety students were selected from grades six, eight, and ten in the P.K. Yonge Laboratory School, Gainesville, Florida. These students had birthdays in one of the following twelve month time periods; beginning in December 1952, 1954, 1956. The experiment was conducted in October 1968, therefore, the sample subjects in grades six, eight, and ten were approximately eleven, thirteen, and fifteen years of age, respectively. A stratified random sampling process placed five students in each of eighteen cells. The cells were formed with respect to three grade levels, three ability levels, and sex. The intelligence scores from the California Test of Mental Maturity served as the basis for ability stratification for grades six and eight. General ability percentiles from the School and College Ability Test (SCAT) served the corresponding role for the tenth grade subjects. The statistical technique employed was a factorial analysis of variance using grade, ability, and sex as independent variables.

A significant difference in favor of the boys existed between the mean scores for each sex on the total sectioning test. This significant difference remained when the test scores were analyzed with respect to performance on each of the four solids and on each of the four types of cuts performed on the solids. Sixth grade students scored significantly below both the eighth and tenth grade students on the total test and on every cut and solid. Low-ability students scored significantly below both middle- and high-ability students on the total test, on the cylinder and on cuts along the major and minor axes of each solid. For the given sample of subjects, the cone was the most difficult solid to section and the oblique cut was the most difficult to visualize.

For the sample studied, age was accepted as a significant factor in the development of the ability to visualize cross-sections. The median statistic for grades eight and ten was interpreted as generally supporting Piaget's position that children of approximately twelve years of age can visualize the conic sections. The medians for grades eight and ten were 13.5 and 14.5 (out of a perfect score of 16). The median for grade six was 11.5.

Using a criterion of a perfect score on her multiple choice test, Boe concluded that the raw data did not substantiate Piaget and Inhelder's statement that children of twelve years of age have achieved mastery of the geometric sections (only three of her seventy-two subjects in grades eight, ten and twelve correctly responded to all sixteen items). It should be observed that the acceptance of age as a factor in the development of the ability to perceive sections in this study is based upon what the researcher felt were acceptable levels of performance (median statistics) and not upon 100% mastery of the test items.

Sex was found to be a significant variable affecting test performance. While Boe found that boys' scores, on her multiple choice test, tended to be higher than girls these differences were not significant at the .05 level.

The results of this study seem to be consistent with Boe's with respect to ability as a significant variable in test performance i.e., cross-sectioning ability is higher for higher ability students.

While more research is needed to help resolve and explain the different findings in relation to sex and the differing conclusions as regard to age, between this study and Boe's, this investigator believes that his results indicate that cross-sections may reasonably be utilized in most mathematics programs provided the students have a prior opportunity to participate in some actual sectioning experiences.

Research Reporting Section 2

Leader: Kenneth J. Travers, University of Illinois, Urbana, Illinois.

Speakers: 1. Larry C. Elbrink and Bert K. Waits, The Ohio State University, Columbus, Ohio, "Student Evaluation of the Teacher in the Mathematics Classroom - Is it Meaningful."

2. Jane Swafford and Len Pikaart, University of Georgia, Athens, Georgia, "The Relationship Between Personality and Differential Achievement in Eighth Grade Mathematics."

3. Ben V. Flora, Jr., Northern Illinois University, Dekalb, Illinois, "Diagnosing Teacher Behavior Characteristics of Teachers of Secondary School Mathematics."

4. Joseph W. Dodson, Western Carolina University, Cullowhee, North Carolina, "A Study to Determine Characteristics of Students Who are Successful in Solving Insightful Mathematics Problems."

STUDENT EVALUATION OF THE TEACHER IN THE MATHEMATICS CLASSROOM - IS IT MEANINGFUL

Larry C. Elbrink and Bert K. Waits
The Ohio State University
Columbus, Ohio

The purpose of this study was to test the hypotheses that student evaluations of their (i) teacher, (ii) course, and (iii) examinations, are unbiased relative to individual student mathematics achievement in the course.

Many teachers feel student evaluations are biased toward "liking" the teacher if they are doing "well" in the course and "disliking" the teacher if they are doing "poor" work in the course. Consequently, such teachers often feel student evaluations are of no use (at least to them) and feel "threatened" if the results were to be available to others for interpretation.

The mathematics teacher is particularly prone to feel "threatened" by student evaluations because, by its very nature, many students view mathematics as "hard," "unpleasant," etc. If it can be shown (relative to a particular evaluation instrument) that students do give unbiased (with respect to course achievement) evaluations, then the classroom teacher could view student evaluations as (possibly) beneficial and meaningful.

An evaluation instrument developed by Dr. Robert W. Ullman, Director of the Office of Evaluation at the Ohio State University was used in the experiment. His instrument is in wide use at many universities. It consisted of 48 questions divided into three categories - course, instructor and examinations. There was approximately a 50-50 split between positively phrased questions and negatively phrased questions. Students were given the choice of four responses - strongly agree, agree, disagree, and strongly disagree.

The instrument was given to all students enrolled in Mathematics 117 near the end of the course in the spring quarter, 1969 (Mathematics 117 is the second course in a calculus with economic applications sequence offered at the Ohio State University). There were sixteen individual sections of the course taught by thirteen different teachers (advanced graduate students in the department). Each section instructor followed a common syllabus and all students took common departmental examinations given in the evening.

The students were asked to code in their average midterm score in addition to the other information requested on the Ullman instrument. The forms were processed through a digitek 100 optical scanning system

(IBM card output). A computer program was written that assigned a number (percent) - representing the evaluation rating - to each category (course, instructor, and examinations) by the following formula,

$R = \frac{\sum r_i}{3 \cdot N}$; where N is the total number of questions in the category

and $r_i = 3$ if the response to the i^{th} question is "strongly positive," or $r_i = 2$ if the response to the i^{th} question is "positive," or $r_i = 1$ if the response to the i^{th} question is "negative," or $r_i = 0$ if the response to the i^{th} question is "strongly negative." ("Strongly positive" equals a "strongly agree" response to a "positive" question while "strongly positive" equals a "strongly disagree" response to a "negative" question. Similarly for "positive," "negative," and "strongly negative.")

The following data were gathered for each student at the conclusion of the experiment - student number, section number (1 through 16), course evaluation rating, instructor evaluation rating, examination evaluation rating and average midterm score.

Statistics were computed for each section correlating achievement (average midterm score) with (i) the course evaluation rating, (ii) the instructor evaluation rating, and (iii) the examinations evaluation rating. The following hypotheses were tested (with respect to each section);

(i) H_1 : Course evaluation ratings are independent of student achievement.

(technically: the correlation coefficient between the two variables is zero)

(ii) H_2 : The instructor ratings are independent of student achievement.

(iii) H_3 : The examination ratings are independent of student achievement.

In addition, the students were identified as belonging to one of five "treatment groups" defined as follows:

rank A - average midterm score 90 or above (100 possible),

rank B - average midterm score between 79 and 90,

rank C - " " " " 69 " 80,

rank D - " " " " 59 " 70,

rank E - average midterm score less than 60.

A one-way analysis of variance model was applied with the course (instructor, examination) rating as the dependent variable and the effect for the "treatment" (defined above) as the independent variable. Again, relative

to each section, the following hypotheses were tested:

(i) H_4 : There is no effect for achievement rank on the course evaluation ratings,

(ii) H_5 : There is no effect for achievement rank on the instructor evaluation ratings,

(iii) H_6 : There is no effect for achievement rank on the examination evaluating ratings.

It is worth noting that the analysis of variance model can detect differences in effects for achievement ranking that a correlation analysis could not. The results of these analyses are summarized in the tables which follow.

It is clear (after considering the results summarized in Table Two) that in four sections (numbers 1, 6, 7, and 15) the students gave biased evaluations of their teacher (the higher evaluation ratings were associated with the "good" students while the lower ratings were associated with the "poor" students). However, students in twelve sections have their teachers seemingly unbiased evaluations relative to their achievement in the course.

Students in almost one-half of the sections gave seemingly biased evaluations of the course and examinations. Again the better student gave the more favorable evaluation rating in both categories.

What does this all mean? Without discussing the controversial and much debated problems of intrepretation, value and use of student evaluation results, we can, somewhat surprisingly, conclude that many teachers might expect to receive an unbiased evaluation from their students.

We cannot recommend any evaluation instrument (including the Ullman instrument) nor can we recommend that instructors or administrators encourage student evaluations without a careful investigation of psychological and "local" considerations. It is possible that inexperienced teachers might direct their teaching activities toward developing "favorable" evaluation ratings. Such activities might not be in the student's best interest and they even possibly could result in ineffective teaching. We do recommend that an instructor who uses the Ullman instrument perform a correlation analysis or regression analysis on the variables (evaluation rating and course performance) and then interpret the results accordingly.

Table One

Course Evaluation

Section	N	Correlation coefficient	Reject Hypothesis H_1 (.05 level)	F	Reject Hypothesis H_4 (.05 level)
1.	18	.6556	Yes	7.5477	Yes
2.	21	.0881	No	0.8270	No
3.	25	.0888	No	0.7388	No
4.	20	.5184	Yes	1.6775	No
5.	17	.4368	Yes	1.8655	No
6.	25	.4685	Yes	3.5777	Yes
7.	15	.3718	No	1.8284	No
8.	15	.5324	Yes	2.3282	No
9.	22	-.0013	No	0.3160	No
10.	16	.3660	No	1.0853	No
11.	22	.4588	Yes	1.2698	No
12.	30	.3787	Yes	1.4766	No
13.	17	.0031	No	0.8956	No
14.	17	.1813	No	1.2148	No
15.	21	.5743	Yes	1.9105	No
16.	24	.1734	No	1.0117	No

Table Two

Teacher Evaluation

Section	N	Correlation coefficient	Reject Hypothesis H_2 (.05 level)	F	Reject Hypothesis H_5 (.05 level)
1.	18	.5989	Yes	5.7424	Yes
2.	21	-.1225	No	0.8381	No
3.	25	-.0936	No	1.6780	No
4.	20	.3127	No	0.5119	No
5.	17	.3020	No	0.7033	No
6.	25	.3586	Yes	2.0142	No
7.	15	.5042	Yes	2.1399	No
8.	15	.3284	No	0.3465	No
9.	22	.1013	No	1.7996	No
10.	16	.0160	No	0.2982	No
11.	22	.0801	No	0.1997	No
12.	30	.2142	No	1.0632	No
13.	17	.3801	No	1.1310	No
14.	17	.3514	No	1.5377	No
15.	21	.4236	Yes	2.3657	No
16.	24	.0848	No	2.2916	No

Table Three

Examination Evaluation

Section	N	Correlation coefficient	Reject Hypothesis H_3 (.05 level)	F	Reject Hypothesis H_6 (.05 level)
1.	18	.6091	Yes	2.0520	No
2.	21	.3099	No	0.7124	No
3.	25	.2455	No	1.2285	No
4.	20	.3470	No	1.5865	No
5.	17	.6645	Yes	4.6458	Yes
6.	25	.4917	Yes	1.2300	No
7.	15	.0722	No	1.3297	No
8.	15	.3108	No	0.5188	No
9.	22	.2111	No	0.9740	No
10.	16	.5535	Yes	2.1796	No
11.	22	.5321	Yes	3.1690	Yes
12.	30	-.0235	No	0.1278	No
13.	17	.5452	Yes	2.0609	No
14.	17	.3314	No	1.7851	No
15.	21	.1991	No	0.3322	No
16.	24	.1408	No	1.2421	No

THE RELATIONSHIP BETWEEN PERSONALITY AND DIFFERENTIAL ACHIEVEMENT IN EIGHTH GRADE MATHEMATICS

Jane Swafford and Len Pikaart
University of Georgia
Athens, Georgia

The purpose of the study was twofold. It sought first to investigate achievement in mathematics with respect to (a) computational skills and (b) the understanding of mathematics concepts, and second to determine how achievement in these components of mathematics learning is related to nonintellective personality factors and general mental ability. The study sought to answer these questions. Can test items be selected and categorized a priori which will measure components in mathematics achievement for mathematical computational skill and for understanding of mathematics concepts? If two sets of test items can be developed, what will be the distribution of students classified with respect to their scores on both sets of test items? If different achievement groups are found, can nonintellective personality factors or general mental ability be used to differentiate among these groups?

Little previous research has been done in the empirical determination of the components of mathematics achievement frequently assumed to exist. Mathematics achievement tests which report to measure different components of achievement rely almost exclusively on content validity to support their contentions. In most previous studies investigating the relationship between personality and achievement in mathematics, the latter is viewed as a unitary trait. In most cases, a healthier personality, less anxiety, or extraversion was found associated with better mathematics achievement. The emerging role of personality in mathematics achievement must be refined in light of a multiple components view of mathematics learning.

Test items, selected primarily from the NLSMA eighth grade test batteries, were classified by a panel of mathematics educators using a modification of the hierarchy of cognitive behaviors outlined in the Taxonomy of Educational Objectives, edited by Benjamin Bloom. A sample of 94 of these items was administered to the eighth grade students at a selected, suburban Atlanta, secondary school. Through factor analysis of the scores obtained, two factors, which accounted for 50 per cent of the total variance, were identified. From the factor loadings and the a priori classification of the test items, it was confirmed that the test measured two components of mathematics achievement which could be labeled computational skill and mathematics understanding. Students were then classified as scoring below or above the median in each factor and grouped according to their performance into a low skills--low understanding, low skill--high understanding, high skill--low understanding, or high skill--high understanding category. The comparative size of the

achievement groups was analyzed by computing chi square. Measures of personality and general mental ability were obtained for 335 students using Cattell's Jr.-Sr. High School Personality Questionnaire and the Otis Quick-Scoring Mental Ability Tests, New Edition, Beta Test. Discriminant analysis was used to determine if nonintellective personality factors could discriminate among the resulting groups of mathematics achievers. Analysis of variance and the Duncan's multiple range test were used to examine the differences among the groups on scores on each of the 13 nonintellective, first order, personality factors, the two second order personality factors of extraversion and anxiety, and general mental ability. Differences were accepted as significant at the .05 level.

The hypothesis of equality of group size for the four differential mathematics achievement groups was rejected at the .01 level. The larger groups were the two in which students either scored high on both mathematical computational skill and mathematics understanding or low on both factors. The two smaller groups accounted for nearly 30 per cent of the total population, split evenly between them. According to the theory of the hierarchy of cognitive behaviors, this distribution would be unexpected.

In the discriminant analysis, the first two discriminant functions accounted for a major portion of the discriminating power of the variables. The first function separated the groups whose performance in computational skill and mathematics understanding was consistent. The second function tended to separate the high understanding groups from the low understanding groups.

The analysis of variance showed significant differences among the groups means on 4 of the 13 nonintellective personality factors. The significant differences found between pairs of means, as revealed by Duncan's multiple range test, were more numerous. The personality factors of cyclothymia (a), Ego Strength (C), Excitability (D), Super Ego Strength (G), Self-Sufficiency (O_2), and High Ergic Tension (O_4) accounted for most of the differences observed between the groups. These results indicated that personality differences tended to separate the high skill-high understanding group from the other three groups. However, this group differed significantly from the low skill-low understanding group only on the measure of conscientiousness.

The null hypothesis for the two second order personality factors, anxiety and extraversion, was accepted. However, multiple comparison indicated a significant difference in mean anxiety score between the two high skill groups, the more anxious group having the lower scores in mathematics understanding. A significant difference in extraversion scores between the two high understanding groups was found. The more extraverted group was the group which was also high in computational skill.

Among the four achievement groups, the null hypothesis on measures of general mental ability was rejected. Multiple comparison,

however, revealed that the two groups whose performance in computational skills and mathematics understanding was inconsistent did not differ significantly on measures of general mental ability.

Thus, the study resulted in the development of an instrument that measured two components of mathematics achievement, skills and understandings, and in the identification of four groups of mathematics achievers in reference to these components. These groups differed on measures of intellectual ability; but more interesting, these groups showed measurable differences on nonintellective variables.

DIAGNOSING TEACHER BEHAVIOR CHARACTERISTICS OF TEACHERS OF SECONDARY SCHOOL MATHEMATICS

Ben V. Flora, Jr.
Northern Illinois University
Dekalb, Illinois

In recent years much attention has been given to questions concerning the classroom behavior of teachers. Observing teachers in the classroom, recording a variety of types of information, and carefully subjecting this recorded data to analysis has been a common pattern. During an earlier era, research devoted to teaching or teachers was often focused on the determination of characteristics which were common for effective or ineffective teachers. Thus, the research focus has seemingly changed from one concerned with defining characteristics of teachers to one concerned with studying the classroom behavior of teachers. It is the case, however, that the present attention to behavior in the classroom is often constructed on a base developed by the earlier "characteristics" research. Like many of today's studies, the one reported employs both a "characteristics" and a "behavior" base. This research was directed to the development of and exploratory studies with the Teaching Situation Reaction Test for Teachers of Secondary School Mathematics (TSRT-TSSM). This paper-pencil instrument, which is designed to provide a measure of selected teacher behavior characteristics of teachers of secondary school mathematics, was primarily developed to be used diagnostically in methods courses designed for prospective teachers of secondary school mathematics.

Based upon the assumption that classroom behavior is a function of (among other factors) characteristics or beliefs possessed by teachers, and using results of previous efforts devoted to determining characteristics of effective and ineffective teachers, the researcher selected ten teacher behavior characteristic dimensions for study. A fifty item test which was designed to measure these teacher behavior characteristics was developed by using a "teaching situation approach." The test defines a teaching assignment which places the testee in a junior-senior high school. A number of "teaching situations" are hypothesized to occur during the course of a school year. Each of the fifty items which relate to actions regarding the situations is designed to measure one of the ten teacher behavior characteristics defined for the study. The ten scores produced (one for each of the ten characteristics measured) provide a profile which is purported to indicate the beliefs of an individual concerning these characteristics. This profile for an individual teacher, when compared with profiles representative of highly effective and minimally effective teachers of secondary school mathematics, provides the teacher with information which he may use as he considers the desirability of producing a change in his teacher behavior characteristics.

In order to obtain information concerning the behavior of the test, it was administered to a group consisting of experienced teachers, each one classified as a highly effective or minimally effective teacher of secondary school mathematics, a second group of inservice teachers, preservice teachers with no teaching experience who were enrolled in a first methods course for prospective teachers of secondary mathematics, and preservice teachers with limited teaching experience afforded by student teaching. From an analysis of the results of these testings, information concerning (1) "norm" scores for different types of groups, (2) reliability, (3) resistancy to faking, and (4) relationships between the instrument or its parts and selected variables was obtained.

Making use of portions of the information obtained in these initial testings, the test was administered to members of a methods class at the second class meeting. Five weeks later, the class was provided with information about their test profiles and other results which had been obtained in exploratory work. Each student had information concerning his teacher behavior characteristics and a basis for comparing the measure of his characteristics with those of teachers of mathematics which are classified as highly effective or minimally effective teachers as well as "norm" scores from student teachers and another methods group. Several weeks after the discussion of the test and the characteristics it is purported to measure, the test was readministered and significant changes in scores relating to five of the ten teacher behavior characteristics occurred.

As a result of the exploratory studies, it was possible to conclude that test scores for highly effective teachers were significantly higher than those for minimally effective teachers of secondary school mathematics and that for a number of the teacher behavior characteristics being measured, profiles of the two groups were significantly different. Satisfactory test-retest reliability and an indication that the instrument is resistant to faking was also obtained. Additional results suggest that the profiles of individuals provide a reasonably accurate description of the individual with regard to the characteristics measured and that the instrument can be expected to predict the degree of success in a first mathematics methods course and in student teaching in mathematics.

A STUDY TO DETERMINE CHARACTERISTICS OF STUDENTS WHO ARE SUCCESSFUL IN SOLVING INSIGHTFUL MATHEMATICS PROBLEMS

Joseph W. Dodson
Western Carolina University
Cullowhee, North Carolina

The purpose of the study was to describe successful "insightful" mathematics problem-solvers in terms of: (1) their performance on mathematics achievement tests, (2) their possession of certain cognitive and personality traits, (3) the characteristics of their teachers, and (4) the characteristics of their school and communities. A unique feature of the study was the type of problem used to measure the criterion variable. The term "insightful" was used to describe these problems to emphasize that they have been judged to be challenging and unlike problems which the subjects have likely solved before, i.e., the routines followed in solving "textbook" type problems are not sufficient to produce a correct response.

The ability to solve non-routine mathematics problems which are challenging, but not impossible, for the problem-solver has long been considered an important ability for mathematics students. Currently there appears to be very little known about this ability. This study was designed to acquire knowledge of the student characteristics associated with this ability. Such knowledge could be used by teachers to identify students with potentially high ability to solve insightful problems and by curriculum specialists interested in developing this ability.

Very little previous research has dealt with the ability to solve non-routine or insightful, mathematics problems. Instead, the criterion tests have been measured with instruments containing routine "story" problems similar to those found in elementary or secondary textbooks. The present investigation has tested seventy-six variables--student, teacher, school, and community--for their relevance to problem-solving ability and has determined which of these variables have the strongest relationship to problem-solving ability. Successful problem-solvers are thus characterized on the basis of the presence or absence of the traits or abilities described by the seventy-six variables.

The study was designed to make use of the National Longitudinal Study of Mathematical Abilities (NLSMA) data banks from which all of the necessary data were obtained. This data included the seventy-six measures for each subject in a 10 percent sample (approximatley 1500 subjects) of the NLSMA Z-Population (studied from grade ten through grade twelve) and the subject's response to each of the forty insightful mathematics problems which were selected from the NLSMA test batteries.

Statistical analyses were performed to determine: (1) which variables discriminated significantly among six ability groups designated on the basis of the subject's performance on the criterion test, and (2) the relative strength of the variables as discriminators. Analysis of variance and discriminant analysis were selected as the appropriate statistical models: (1) analysis of variance providing an "over-all" test of significance for the difference among the means of the six ability groups and (2) the discriminant analysis providing a method of comparing the relative strength of the variables as discriminators. The analyses resulted in an ordering of the variables from the best to the poorest discriminators. Comparisons were then made between the test items measuring the best and the poorest discriminators (from among the mathematics achievement variables) to determine the differences in their nature.

Analyses similar to those performed for the total criterion test were performed for the three subtests of the criterion test--the algebra, geometry, and number subtests--to determine whether or not a variable was discriminating among ability groups due to its strong relationship to only one subtest as opposed to a strong relationship to all three, and hence, a strong relationship to the total criterion test.

All of the mathematics achievement variables were significant discriminators among ability groups. The items measuring the weaker discriminators required little synthesis (i.e., the problem required organization of very few mathematical ideas to produce a solution) and the mathematical ideas involved were relatively elementary. The items measuring the strongest discriminators can be described as those which require a great deal of synthesis of sophisticated and/or seemingly unrelated mathematical ideas or as items requiring the subject to solve routine algebraic equations.

Assuming that insightful mathematics problem-solving ability is important, then the development of this ability should include:

a) student exposure to advanced topics in mathematics including the algebra of inequalities involving the solution of quadratic inequalities, or systems of inequalities which involves a considerable amount of synthesis.

b) considerable emphasis on solving geometry problems which require the students to synthesize a large number of seemingly unrelated geometric ideas as opposed to problems solved by the simple application of the Pythagorean or other familiar theorems.

c) emphasis in solving routine algebraic equations to provide the necessary "tools" for solving problems.

d) the study of more elementary mathematical skills and content typical of the items measuring weaker discriminators. However, mathematical study limited to the acquisition of more basic ideas and skills is not likely to produce proficient problem-solvers.

The following is a composite list of the strongest characteristics--aside from superiority in mathematics--of a successful problem-solver. He:

1. scored high on verbal and general reasoning tests.
2. was good at determing spacial relationships.
3. was able to resist distraction, identify critical elements, and remain independent of irrelevant elements.
4. was a divergent thinker.
5. had low debilitating test anxiety; high facilitating anxiety.
6. had a positive attitude toward mathematics.
7. could have been messy or not messy.
8. had teachers with the most credits beyond the bachelor's degree.
9. had teachers with the highest degrees.
10. came from a high-income family.
11. lived in a community with higher starting teacher's salary than his poorer problem-solving peers.
12. lived in a community having a recent population size change.
13. had about the same socio-economic index as a poorer problem-solver.

Research Reporting Section 3

Leader: Joseph Payne, University of Michigan, Ann Arbor, Michigan.

Speakers: 1. James M. Sherrill, University of Texas, Austin, Texas, "The Effects of Differing Presentations of Mathematical Word Problems Upon Student Achievement."

2. Zalman Usiskin, University of Chicago, Chicago, Illinois, "The Effects of Teaching Euclidean Geometry Via Transformations on Student Achievement and Attitudes in Tenth-Grade Geometry."

3. Jack L. Shepler, Indiana University of Pennsylvania, Indiana, Pennsylvania, "The Employment of a Developmental Model to Build a Probability Unit for Sixth Grade Students and Its Effect on Mastery Learning and Retention."

4. Larry Sowder, Northern Illinois University, Dekalb, Illinois, "Performance on Some Discovery Tasks In Grades 4-7."

THE EFFECTS OF DIFFERING PRESENTATIONS OF MATHEMATICAL WORD PROBLEMS UPON STUDENT ACHIEVEMENT

James M. Sherrill
The University of Texas
Austin, Texas

The ability of students to solve mathematical word problems has long interested mathematics educators. In 1945 The Commission on Post-War Plans listed the solving of verbal word problems as essential in the mathematics curriculum of the public schools. Kilpatrick indicated that "Problem solving. . .(has) received increased attention from mathematics educators in the past five years." (3: 1) The continuing interest in problem solving in mathematics is, in part, a product of the importance of problem solving in the mathematics curriculum. The present study's interest in problem solving is limited to printed mathematical word problems. The members of the Cambridge Conference listed, in the Goals for School Mathematics, four reasons why work with mathematics word problems should be included in the mathematics curriculum. The report states that mathematics word problems should:

1. illustrate and reinforce the ideas in the corresponding portion of the text

2. provide continual review

3. furnish practice in computation

4. furnish the student with good reason for wanting to know the answers to arithmetical problems (2: 28)

The method of presentation of printed mathematical word problems was one of the independent variables in the present study. The use of a pictorial representation of the problem situation was the criterion which determined the different methods. When the study was being planned there were two groups--those who did and those who did not have a pictorial representation of the problem situation. Out of the process of reviewing the literature a third group evolved. Trimble, who was considered as being in favor of using a pictorial representation of the problem situation, pointed out that when constructing the pictorial representation one could include "errors made on purpose". Trimble felt that such errors could be used as ". . .natural, built-in graphical checks; and you can expect reactions like 'No, that couldn't be right!' once the boys and girls become personally involved in the solution of the problem." (5: 7)

Brownell, in his general discussion of problem solving, offered a list of twelve practical suggestions for developing ability in problem solving. Suggestion j) was as follows:

> Part of the real expertness in problem solving is the ability to differentiate between the reasonable and the absurd, the logical and the illogical. Instead of being

> 'protected' from error, the child should many times be exposed to error and be encouraged to detect and to demonstrate what is wrong, and why. (1: 440)

The purpose of the present study is to test the following hypotheses:

> Presenting a pictorial representation with the problem situation of a mathematical word problem has no effect upon student achievement, independent of the accuracy of the pictorial representation.
>
> The subject's I.Q., reading score, and grade point average in the preceding year's mathematics courses will have no effect upon his achievement.

The test was constructed by the author, but the problems included in the testing instrument were from two main sources: (a) the Y- and Z-Population test batteries of NLSMA and (b) three mathematics textbooks currently being used at the tenth grade level. The criterion for inclusion of a problem was that it could be solved with or without a pictorial representation. Two tests were constructed from the problems selected and each test was given to a group of 21 students taking the required course for elementary education majors at The University of Texas at Austin. The scores of the first and fourth quarters of the students, as determined by their total test score, were then compared for each of the problems. The twenty problems that discriminated the best were selected for inclusion in the final testing instrument. There were three forms of the instrument: Form A--consisted of the twenty selected mathematical word problems; Form B--contained the same information as Form A with the addition of an accurate pictoral representation of the problem situation; Form C--contained the same information as Form B except the pictorial representation had scale distortions and/or some misrepresented relationships. Since there was simply a need for an accurate and a distorted pictorial representation there was no work done to determine whether one way of distorting a pictorial representation was "better" than another way of distorting the same pictorial representation. Student achievement was operationally defined to be a subject's score on the form of the testing instrument he took.

Three groups of tenth graders were constructed--Group A was administered Form A; Group B was administered Form B; Group C was administered Form C. Since the school system would not let the researcher randomly assign subjects to the three groups he randomly ordered the three forms of the test. The tests were administered in the student's normal mathematics classroom. The randomly equated groups consisted of 322 tenth graders. Data collected from the school records were the subject's I.Q. score, reading score, and his grade average in mathematics in the ninth grade. Unfortunately the D. A. T. scores were not yet available.

The means of the three groups were compared by single classification of analysis of variance. The means were then ranked according to Kramer's work with the Duncan Multiple Range Test for unequal N. (4: 307-309) An item analysis was run and a Cronbach alpha of .7007 was attained. The data for each group was

also divided according to I.Q. score, reading score, and grade point average in mathematics in the ninth grade. The means were then compared by double classification of analysis of variance.

The comparison of the means of the three groups yielded an F of 80.623 ($P < .0001$) so the means were ranked. Application of the work of Kramer ranked the three means as follows: (1) Group B (2) Group A (3) Group C. There was a significant difference between each pair of the three means. Each of the double classification analyses yielded significant F ratios for both main effects and the interaction effect. In every case the probability level was less than .01.

Support was found for rejection of the first hypothesis. The method of presenting the mathematical word problems did have an effect upon student achievement. Presenting a mathematical word problem with an accurate pictorial representation was most beneficial. Presenting a distorted pictorial representation of the problem situation hinders student achievement. The second hypothesis also faces the chance for rejection since, in this study, all three (I.Q. score, reading score, and grade point average in mathematics in the ninth grade) variables had an effect upon the subject's achievement.

BIBLIOGRAPHY

1. Brownell, William A. "Problem Solving" The Psychology of Learning. The Forty-first Yearbook of The National Society for the Study of Education. Chicago: The University of Chicago Press, 1951.

2. Cambridge Conference. Goals for School Mathematics. Boston: Houghton Mifflin Company, 1963.

3. Kilpatrick, Jeremy. "Problem Solving and Creative Behavior in Mathematics" Draft of a chapter of the October, 1969 Review of Educational Research.

4. Kramer, C. Y. "Extension of Multiple Range Tests to Group Means With Unequal Numbers of Replications" Biometrics, 1956, 12, 307-309.

5. Trimble, Harold C. "Problems as Means" The Mathematics Teacher. Washington D. C.: The National Council of Teachers of Mathematics, Vol. LIX, number 1 (January 1966).

THE EFFECTS OF TEACHING EUCLIDEAN GEOMETRY VIA TRANSFORMATIONS ON STUDENT ACHIEVEMENT AND ATTITUDES IN TENTH-GRADE GEOMETRY[1]

Zalman Usiskin
University of Chicago
Chicago, Illinois

Prominent mathematics educators (e.g., Adler (1968)[2] and Allendoerfer (1969)[3]) have suggested that the study of geometric transformations be one of the major goals of tenth-grade geometry. No earlier researcher seems to have studied the feasibility or possible effects of employing point transformations as a fundamental concept in a tenth-grade Euclidean geometry course.[4]

This study was conducted during the school year 1968-69 using materials developed and taught by Coxford and Usiskin[5] in the previous school year. These materials used (1) preservation properties of reflections as postulates, (2) definitions of congruence and similarity in terms of transformations, (3) reflection proofs to introduce the student to his own proofs of statements, (4) proofs of all triangle congruence and similarity propositions by means of transformations, (5) matrices to represent transformations. These features served to distinguish the experimental materials from other contemporary geometry texts in which transformations are seldom (if ever) used and almost never integrated into the mathematical development of the course.

This study sought to shed light upon four major questions:

1. What is the effect of the experimental materials on student learning of geometry concepts considered as standard material in contemporary programs?

2. What is the effect of the experimental materials on student attitudes towards mathematics?

3. Can students attain competence with the mathematical concepts unique to this transformation approach?

4. Can the experimental materials be implemented with little or no external guidance?

Other minor questions were studied; answers to these are indicated later in this article.

Experimental and control populations were employed. The experimental population consisted of approximately 425 students in 6 schools, taught by 8 teachers. Students of all abilities in communities of various sizes were represented. The teachers were primarily volunteers, having heard of the study from talks given by the experimenter. A similarly diverse control population, with 475 students, 9 teachers, and 7 schools, was located with the aid of personal contacts. Control classes tended to be more heterogeneous and larger than experimental classes.

At the beginning of the school year, compared to experimental students, control students had (1) significantly greater knowledge of standard geometry content as measured by the ETS Cooperative Geometry Tests, Part I, Forms A and B, (2) greater skill with algebra and arithmetic, and greater ability to perceive aspects of figures and motion, as measured by an experimenter-constructed instrument, (3) approximately similar attitudes, as measured by the Aiken-Dreger[6] instrument, and (4) slightly lower socio-economic levels as measured by a scale of Warner et al[7]. Control teachers had nearly the same backgrounds in teaching and courses as experimental teachers.

A variety of texts was used by control classes, quite possibly reflecting nation-wide tendencies. These classes generally devoted greater lengths of time to topics common to the experimental and control texts. Experimental classes devoted approximately 7 weeks of class time (interspersed throughout the year) to material unique to the experimental approach, not including time spent on standard topics which utilized transformations in their development.

The instruments used at the beginning of the year were also used at the end of the year, each student receiving a form of the test different from the form taken in September. In addition, a test of content unique to the experimental approach was given to the experimental students. The attitude scale was given four times during the year: in September, December, March and June.

At the end of the school year, on the ETS Geometry Tests, mean scores of the control population were significantly higher (at the .01 level) than those of the experimental population. These results held even when adjustments were made (in analysis of covariance) for the September scores. This trend held for students of all ability levels, for each sex, and for each part of the ETS examinations, though not always significantly. Experimental students rather consistently correctly answered about 94% as many questions as control students.

Attitudes of both experimental and control students were lower in June than in September, significantly lower for each sex subpopulation except experimental males. Comparisons consistently favored the experimental populations, occasionally at significant levels.

Performance of experimental students on the test of content unique to the experimental approach seemed to indicate a level of comprehension equal to the level of comprehension of standard geometry content.

Opportunities for aid and consultation were offered to experimental and control teachers but were not needed or accepted by either group.

Student ability to perceive aspects of figures and motion improved from September to June, but no consistent pattern favored either experimental or control students, in spite of the fact that certain questions were probably biased in favor of the experimental students.

Experimental students generally outperformed control students on the tests of algebraic and arithmetic skills, but not at significant levels.

A greater percentage of experimental students planned to take mathematics the following year. Experimental and control students seem to have experienced nearly equal ease (or difficulty) in understanding their textbook without teacher explanation, in spite of reports from experimental teachers that their text was notably easy to read.

Relative to the major questions asked, this experimenter has concluded:

1. Students using the experimental materials suffer in their ability to solve types of problems considered standard in contemporary courses.

2. The experimental materials do not have as harmful an effect on student attitudes as other contemporary materials.

3. Students can attain competence with the mathematical concepts unique to the transformation approach.

4. For these materials, at least, teachers do not seem to require retraining for implementation.

One also concludes that it is possible to develop a geometry course based upon transformations for the average student. This forces consideration of the question: Is it worth studying a geometry course utilizing transformation ideas at the expense of time which might be spent in solidifying knowledge of more standard geometry concepts? Answers to this question determine the type of future research which should be undertaken. This study seems to show that the geometry curriculum is at least faced with a choice.

REFERENCES

1. This study is detailed in a Ph.D. dissertation with the same title, completed at The University of Michigan, 1969.

2. Irving Adler, "What Shall We Teach in High School Geometry?", The Mathematics Teacher, 61:226-38 (May, 1968).

3. Carl B. Allendoerfer, "The Dilemma in Geometry," The Mathematics Teacher, 62:165-69 (March, 1969).

4. Related research has been done by Sair Ali Shah ("Selected Geometric Concepts Taught to Children Ages Seven to Eleven," The Arithmetic Teacher, 16:119-28 (February, 1969)) and Earl Beard (An Axiom System for High School Geometry Based on Isometries, Ph.D. Thesis, University of Wisconsin, 1968). There is also some ongoing research in this area being conducted by Olson at the University of Wisconsin and Kort at Northwestern University. The latter includes a follow-up of part of this study.

5. Arthur F. Coxford and Zalman P. Usiskin, Geometry-A Transformation Approach (Ann Arbor, Michigan, the authors, 1968) These materials are presently (1969-70) being used in 25 schools with over 2,000 students.

6. Lewis R. Aiken, "Personality Correlates of Attitude Towards Mathematics," Journal of Educational Research, 56:476-80 (May, 1963).

7. Warner, Meeker, and Eels, Social Class in America (New York: Harper and Roe, 1960).

THE EMPLOYMENT OF A DEVELOPMENTAL MODEL TO BUILD A PROBABILITY UNIT FOR SIXTH GRADE STUDENTS AND ITS EFFECT ON MASTERY LEARNING AND RETENTION

Dr. Jack L. Shepler
Indiana University of Pennsylvania
Indiana, Pennsylvania

Probability and statistics are important mathematical tools used by man in technological society. For numerous reasons, recommendations have been made for a comprehensive program in probability and statistics which begins in the elementary school. Research seems to indicate certain topics in probability and statistics may be suitable for elementary students to learn. However, very few schools are teaching these concepts.

The purpose of this study was three fold: (1) to test the feasibility of teaching topics in probability and statistics to a class of sixth-grade students; (2) to construct a set of instructional materials and procedures in probability and statistics for sixth-grade students; and (3) to investigate the effect on retention of mastery learning of an objective.

The study used the working paper of Shepler, Harvey, and Romberg (1969) and the developmental model of Romberg and DeVault (1967) to build the unit. Shepler et al constructed a framework for the development of an instructional system in probability and statistics for use in the elementary school. The present study was designed to test the feasibility of parts of the working paper.

From strands of the task analysis, the author decided upon behavioral objectives for the unit of instruction and the order in which objectives would be taught. Using this basis, an instructional analysis of the unit was undertaken. The purpose of this analysis was to select or develop materials and procedures for teaching the unit of probability to sixth-grade students.

To aid in the developmental processes of task analysis and instructional analysis, a pilot study was conducted in the fall of 1968. The data from the pilot study was used to identify a set of nine lessons that could be formatively evaluated to test the feasibility of the instructional analysis. The lessons were used to teach a class of sixth-grade students of average to above average ability over a four week period.

The basic instructional procedure for the study approached probability concepts in an intuitive fashion where the student was actively involved in using physical models. He gathered empirical data from experiments and interpreted the results. The student empirically validated major objectives.

The measured objectives were concerned with number of outcomes of an event and a sample space; probability of a simple, compound (including "and", and "or" problems), certain and impossible events; order of two fractions; most likely or equally likely events; empirical probability; likely bounds; law of averages; estimate of the probability; and also subjective notions of probability. Problems tackled by students centered mainly on one and two dimensional sample spaces.

The goal of instruction was to demonstrate mastery of learning of the behavioral objectives. The feasibility of the unit was tested by employing a pretest and posttest. A retention test (the same as the pretest and posttest) was administered 30 days after the posttest.

On the basis of the 72 item criterion pretest and posttest results, one can conclude that the instructional treatment was highly successful regarding achievement on 10 of the 14 measured objectives. Instruction was moderately successful for Objective 10 (Law of Averages 21/25 students scored 100% on Objective 10). In percentages, the average pretest score was 37.9 percent and the average posttest score was 92.8 percent. There was a marked change in student behaviors for all the measured objectives. On this basis, the results of the study support the feasibility of teaching most of the included topics in probability and statistics to the group of students used in this study. However, three objectives were not close to meeting the stated criteria. From a close analysis of the items measuring the three objectives, the author concluded the objectives were not achieved due to a lack of practice in a written situation.

The mean percentage on the retention test was 89.5% (92.8% on posttest). Of the twenty-five children, only five of their scores dropped more than 4% from the posttest to the retention test. Eight children's scores remained the same or increased. Two children who were non-masters of many of the objectives of the unit accounted for 2.4% of the 3.3% decrease. Performance on the ten mastered objectives remained quite high. For this group, mastery learning of the objectives resulted in retention far greater than what normally is achieved.

In the opinion of the author, major reasons for the large gain in raw score can be attributed to the developmental analysis used and the mastery learning techniques that were employed. In the author's opinion, developing a curriculum through the following sequence is an excellent way of building research based curriculum materials. Start with a content outline and establish behavioral objectives. Task analyze these objectives and write an instructional treatment to meet them. Proceed to the important step of actually trying these materials with children, while recognizing the possibility of iteration through preceding steps.

The developmental model encourages modifications of material and procedures based on empirical evidence. Modifications are needed in the unit in light of observations and test analyses.

BIBLIOGRAPHY

1. Bloom, Benjamin S. "Learning for Mastery." UCLA CSEIP Evaluation Comment. Los Angeles: Center for the Study of Evaluation of Instructional Programs, Vol. 1, No. 2, 1968.

2. Romberg, Thomas and DeVault, Vere. "Mathematics Curriculum: Needed Research." Journal of Research and Development in Education, Vol. 1, No. 1, Fall, 1967. Athens: The University of Georgia, 1967.

3. Shepler, Jack L., Harvey, John G., Romberg, Thomas A. "A Proposal and Task Analysis for Teaching Probability and Statistics Concepts in the Elementary School." Working Paper from the Research and Development Center for Cognitive Learning, University of Wisconsin, 1969, in preparation.

4. Shepler, Jack L. "A Study of Parts of The Development of a Unit in Probability and Statistics for the Elementary School." Unpublished Ph.D. dissertation, University of Wisconsin, 1969.

PERFORMANCE ON SOME DISCOVERY TASKS IN GRADES 4-7

Larry Sowder
Northern Illinois University
Dekalb, Illinois

The study had two aims: (1) to explore the ability of pupils of Grades 4, 5, 6, and 7 to give operational evidence of generalizing in selected numerical situations, and (2) to study the effects of differing manners of verbalizing a generalization on the retention of the ability to use the generalizations. The first of these objectives is of interest since there is little reported research concerning how much information students seem to need before they form generalizations; data regarding this matter would be of value to teachers interested in discovery learning. The second aim is related to Gertrude Hendrix's hypothesis that verbalizing a newly formed generalization has a negative effect on retention of ability to use the generalization. This hypothesis, of course, is of concern to all teachers.

Three randomly chosen pupils from each of 24 grade-IQ level-sex blocks were tested individually on a test of eight randomly ordered items. Each item consisted of the stimulus portions of instances of a generalization. For example, for the generalization that the sum of the first n odd numbers is n^2, the student was presented with incomplete instances like 1+3+5+7+9=?. After a short time interval, answers to the problems were provided if the student failed to give the correct response. The number of generalizations formed, as evidenced by consecutive correct responses, and the number of instances required before giving consecutive correct responses provided the dependent measures for problem (1).

Immediately after the test, each pupil was treated on the items on which he was successful in one of three ways: (a) he reviewed the items with no verbalization, (b) he was required to give a correct verbalization of his version of the generalization, or (c) the tester verbalized a correct statement of the generalization. A retention test containing instances of the items was given one week later.

The analyses for problem (1) consisted of univariate grade x IQ level x sex analyses of variance of the dependent measures, and univariate multiple linear regression analyses of these measures, with independent variables chronological age, IQ, arithmetic achievement scores, and mathematical interests scores. For problem (2), retention test scores on selected items were to be analyzed by a one-way analysis of variance.

When pupils did form generalizations, grade means from 4 to 5 instances were required. Over all items, means of 6 to 8 instances resulted. Performance in both the number of instances required and the number of generalizations formed seemed to reach a plateau at Grade 6. There were statistically significant (.01) differences only among IQ level effects and among grade effects for both the total number of generalizations and the

total number of instances. A *post hoc* analysis of the grade effects showed significant (.05) differences only between Grade 4 and each of the other grades. On a score which combined instances and number of generalizations, the Grade 5-Grade 6 difference approached significance at the .05 level. The regression equations for these two variables accounted for slightly more than half the respective variances, with age, IQ, and a computation achievement quotient being the most important variables. Retention data indicated no treatment differences although retention was so scanty that no practical significance could have been attached to significant differences, had they appeared.

Although, not surprisingly, pupils of lower IQ require more instances, indications are that most pupils can indeed form generalizations of the type encountered in the study. With the number of instances needed as a criterion, the optimal grade level at which to offer generalizing tasks appears to be Grade 6 or after. The plateau at Grade 6 supports Piagetian thought, although it may be due to a plateau in computational proficiency.

Research Reporting Section 4

Leader: Len Pikaart, University of Georgia, Athens, Georgia.

Speakers: 1. Robert S. Matulis, University of Florida, Gainesville, Florida, "A Survey of the Understandings of Selected Concepts of Logic by 8-18-Year-Old Students."

2. Robert C. Frazier, Sr., Atlantic Christian College, Wilson, North Carolina, "A Comparison of an Implicit and Two Explicit Methods of Teaching Mathematical Proof Via Abstract Groups Using Selected Rules of Logic."

3. Irv King, University of Hawaii, Honolulu, Hawaii, "Proof in the Elementary School - A Feasibility Study."

4. Richard James Shumway, The Ohio State University, Columbus, Ohio, "The Role of Counterexamples in the Development of Mathematical Concepts of Eighth Grade Mathematics Students."

A SURVEY OF THE UNDERSTANDINGS OF SELECTED CONCEPTS OF LOGIC BY 8-18-YEAR-OLD STUDENTS

Robert S. Matulis
University of Florida
Gainesville, Florida

Interest in the relationship between logic and mathematics has increased tremendously in the past decade. Much needs to be learned about the relationship, however. The main purpose of this survey was to determine whether age, intelligence, sex, or socioeconomic status of eight- to eighteen-year-old children made a difference in their understandings of some of the types of logic which are used in mathematics.

A two-part, eighteen-item, multiple-choice test of students' understandings of implications, conjunction, disjunction, and quantifiers was written by the researcher. Its validity was checked by means of (1) a jury of experts in mathematics, education, and/or logic, (2) a pilot study, (3) computer analyses, and (4) the Spache test of reading level. The test was administered by the researcher to forty-six fourth- through twelfth-grade classes in a county school system of about 75,000 students. Scores of 860 students of both sexes and a wide range of I.Q.'s and socioeconomic levels were used in the study. Their scores were coded according to sex, intelligence group (most recent Kuhlman-Anderson I.Q.: 60-89, 91-109, 111-170+), socioeconomic group (low, medium, or high according to the Otis Dudley Duncan Scale), and age (8-10, 11-13, 14-17).

F tests were computed on scores to determine whether age, intelligence, sex, or socioeconomic status made a statistically significant difference in test scores. More F tests were computed to determine whether interactions of significant variables also made a difference in scores. Multiple-range tests and t tests were done when F tests indicated statistically significant differences in the means of groups to locate the specific means which had statistically significant differences. F tests were calculated to find significant differences in the variances of age groups. Tables and graphs were used in presentation of the data. The 0.05 level of significance was used throughout this study.

Some of the conclusions which can be drawn from this study are the following:

1. Age, intelligence, and socioeconomic status are statistically significant factors in students' understandings of deductive logic, but sex is not a statistically significant factor. In general, older students, brighter students, and students of higher socioeconomic status had higher mean scores than students relatively younger, less bright, or of lower socioeconomic status, respectively.

2. An analysis of mean scores of (Age x Intelligence) interaction groups was done. Mean total-test scores were determined for each age for each intelligence group; they are shown graphically in Figure 1. The darkened lines in the graph indicate year-intervals in which there was a statistically significant difference in mean scores. The dotted lines in the graph point out ages at which there were statistically significant differences in means between the indicated intelligence groups.

Several important conclusions can be drawn from Figure 1. The students in the upper-intelligence group attained a higher level of understanding of deductive logic than did those in the middle group, and those in the middle group a higher level of understanding than those in the lower group. There were three year-intervals of statistically significant growth for students in the upper-intelligence group; there were two for students in the middle-intelligence group; and there was one for students in the lower-intelligence group. The differences in mean scores of the three intelligence groups were all statistically significant after age fourteen.

An examination of Figure 1 will reveal remarkably similar "growth" patterns for the two upper-intelligence groups. There was a period of slow increase in understanding of logic followed by a brief period (i.e., one year) of pronounced growth, a period of slow growth, a second one-year period of pronounced growth, and another period of slow growth. Children in the upper-ability group made considerable gains in understanding of logic between age ten and age eleven and also between age thirteen and age fourteen. Children in the middle-ability group, however, made considerable gains a year later in their lives - between age eleven and age twelve and also between age fourteen and age fifteen. Children in the low-ability group made a considerable gain between the ages of sixteen and seventeen. It appears that students showed the greatest improvement in understanding of deductive logic at a <u>mental age</u> between ten and eleven and also at a <u>mental age</u> between fourteen and fifteen. The size of the low-ability group was small for some ages; consequently, the mean scores may not be representative for some groups. The t tests showed that the means of the two upper-intelligence groups were significantly different at almost every age.

3. An analysis of mean scores of (Age x Socioeconomic) interaction groups was done. Mean total-test scores were determined for each age for each socioeconomic group; they are shown graphically in Figure 2. The darkened lines in this graph indicate year-intervals in which there were statistically significant differences in mean scores. The dotted lines in the graph point out ages at which there were statistically significant differences in means between the indicated socioeconomic groups.

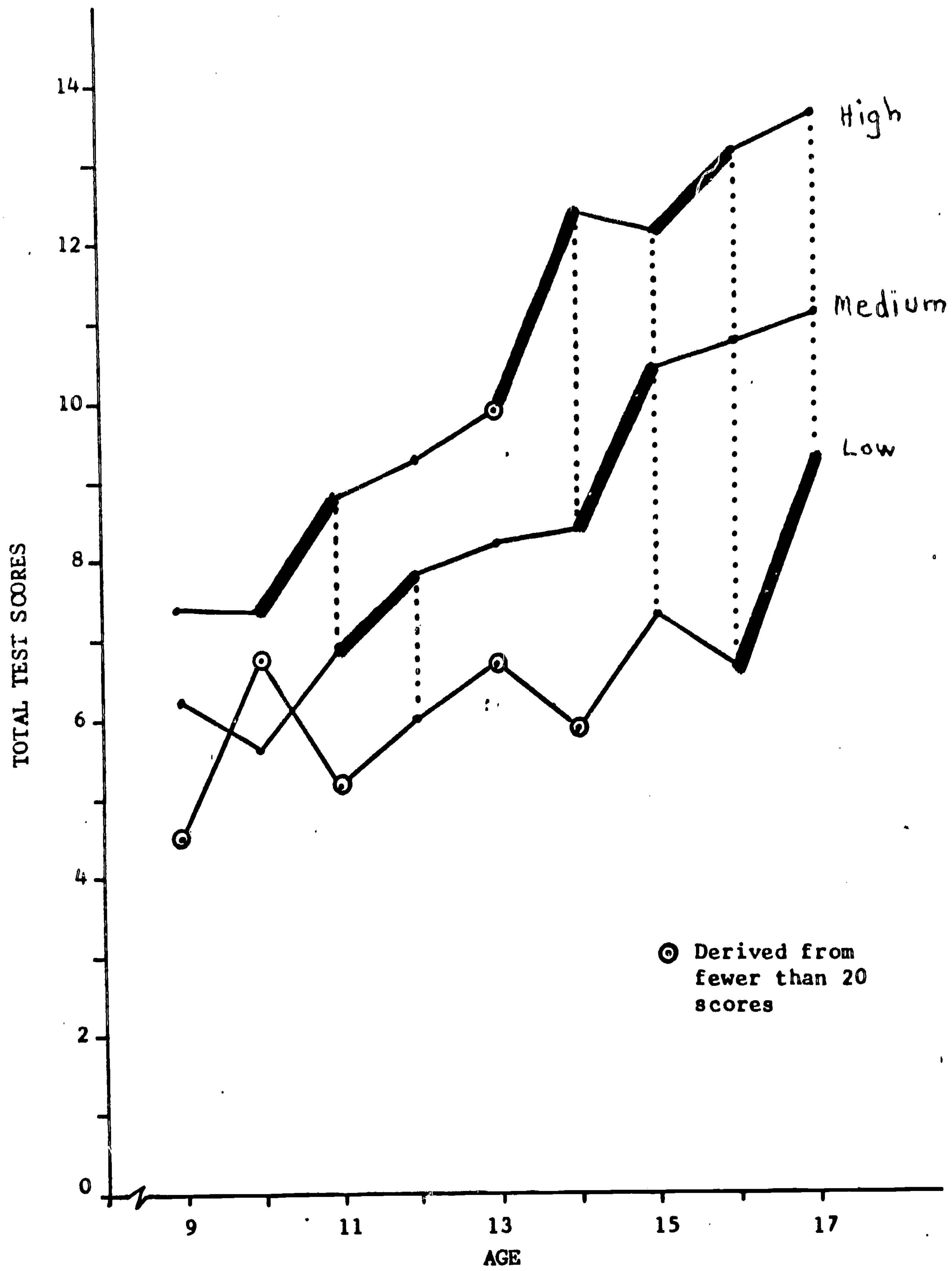

Figure 1. Mean Scores for Each Intelligence Group for Each Age on the Total Test

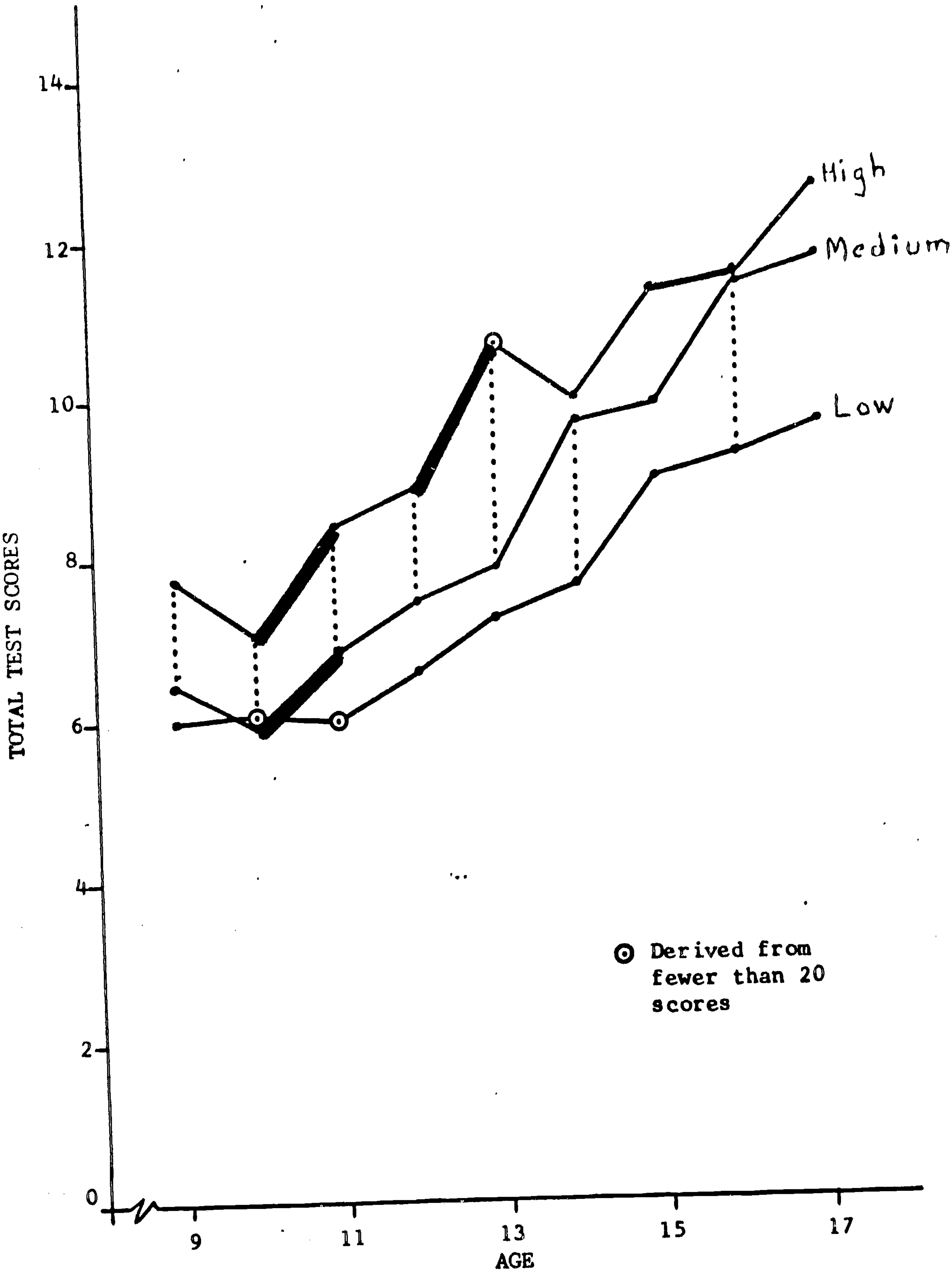

Figure 2. Mean Scores for Each Socioeconomic Group for Each Age on the Total Test

Several important conclusions can be drawn from Figure 2. Those students in the upper-socioeconomic group attained a higher level of understanding of deductive logic than did those in the middle group, and those in the middle group a higher level of understanding than those in the lower group. Each socioeconomic group showed growth in their understandings of deductive logic in almost every year-interval shown. There were two year-intervals of statistically significant "growth" for students in the upper-socioeconomic group; there was one for students in the middle-socioeconomic group; and there was none for students in the lower-socioeconomic group. The period of greatest increase in understandings of logic occurred earlier for students in higher socioeconomic groups. A period of greatest "growth" occurred between ages twelve and thirteen for the high-socioeconomic group, between the ages thirteen and fourteen for the middle-socioeconomic group, and between the ages fourteen and fifteen for the low-socioeconomic group.

The middle-socioeconomic group showed the greatest overall gain in scores from age nine to age seventeen. At ages nine, ten, eleven, twelve, and thirteen, there were statistically significant differences between the means of the middle- and upper-socioeconomic groups; there was also relatively little difference between the means of the middle- and low-socioeconomic groups. At age fourteen and sixteen, however, there was a statistically significant difference between the means of the middle- and low-socioeconomic groups; then there was relatively little difference between the means of the middle- and upper-socioeconomic groups.

4. The variability of scores of fourteen- to eighteen-year-old students was significantly greater than the variability of scores of eleven- to thirteen-year-old students, which in turn was significantly greater than the variability of scores of eight- to ten-year-old students.

Suggestions were made for further research.

A COMPARISON OF AN IMPLICIT AND TWO EXPLICIT METHODS OF TEACHING MATHEMATICAL PROOF VIA ABSTRACT GROUPS USING SELECTED RULES OF LOGIC

Robert C. Frazier, Sr.
Atlantic Christian College
Wilson, North Carolina

The purpose of this study was to investigate experimentally the relative effectiveness of three instructional methods of teaching mathematical proof using selected rules of logic for college freshmen students. The mathematical content of the study was a unit of sixteen elementary abstract group theorems and corollaries.

One instructional method, C, was an implicit method of teaching. Here the rules of logic underlying mathematical proof were neither identified for the student nor designated by name. Rather, the rules of logic were presented solely by their applications in examples and illustrations. The students learned these rules, if at all, only indirectly and/or incidentally from their applications.

Another instructional method, E_2, was an explicit method of teaching. In this method the rules of logic underlying mathematical proof were clearly stated and identified by name. In addition, their uses in mathematical proofs were specified and emphasized at their initial appearance and several times thereafter.

The third instructional method, E_1, was a combination of explicit and implicit teaching. In E_1 the logical content was presented first in an explicit fashion. Then the applications of the rules of logic underlying mathematical proof in elementary abstract group theory were presented in an implicit manner.

Fifty-five students enrolled in a freshman mathematics course in a liberal arts college were randomly assigned to three groups for the two-week study. All groups were pretested by an experimenter-developed test. Each group then studied an experimenter-written program that employed one of the three instructional methods. Two programs, studied by the experimental groups, included at the beginning a unit on selected rules of logic. In the remaining program, studied by the control group, a placebo unit was substituted for the logic unit. All three programs included a development of proofs of the same sixteen theorems of elementary abstract group theory, with one group receiving a treatment that explicitly pointed out some of the uses of rules of logic in the proofs, Upon completion of their respective programs, all subjects were administered an experimenter-developed immediate posttest. Eleven and one-half weeks later all subjects were administered an experimenter-developed delayed posttest.

The 0.05 level of significance was used for the analyses of the subjects' responses to the three tests. These analyses employed analysis of variance techniques for a randomized groups design and yielded the following results.

1. There was a statistically significant difference in means between the groups for section IV of the immediate posttest. Section IV tested for transfer to new materials. The group E_2 receiving the completely explicit treatment performed significantly better than either of the other groups.

TABLE 1.--Analysis of variance of scores on section IV of the immediate posttest.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	*F* ratio
Between	2	70.25	35.13	8.51***
Within	52	214.95	4.13	
Total	54	285.20		

***significant at the 0.001 level

TABLE 2.--Analysis of variance of scores on section IV of the immediate posttest utilizing individual degrees of freedom.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	*F* ratio
Between	2	70.25	35.13	8.51***
Exp. vs. C	1	42.93	42.93	10.39**
E_1 vs. E_2	1	27.32	27.32	6.6[illegible]*
Within	52	214.95	4.13	
Total	54	285.20		

***significant at the 0.001 level
**significant at the 0.01 level
*significant at the 0.05 level

2. There was a statistically significant difference in means indicated by the immediate posttest between group C receiving the implicit treatment and group E_2 receiving the explicit treatment in favor of the explicit treatment.

TABLE 3.--Analysis of variance of scores on the immediate posttest utilizing individual degrees of freedom.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	$\underline{F}$ ratio
Between	2	204.94	102.47	2.70
E_1 vs. 1/2 $(C + E_2)$	1	14.83	14.83	0.39
C vs. E_2	1	190.11	190.11	5.01*
Within	52	1,973.86	37.96	
Total	54	2,178.80		

*significant at the 0.05 level

3. There was a statistically significant difference in means for section II of the delayed posttest that tested for rule of conditional proof and knowledge of elementary abstract group theory. Group E_2 receiving the completely explicit treatment performed significantly better than group C receiving the implicit treatment.

TABLE 4.--Analysis of variance of scores on section II of the delayed posttest utilizing individual degrees of freedom.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	$\underline{F}$ ratio
Between	2	13.14	6.57	2.59
E_1 vs. 1/2 $(C + E_2)$	1	0.92	0.92	0.36
C vs. E_2	1	12.22	12.22	4.81*
Within	46	116.86	2.54	
Total	48	130.00		

*significant at the 0.05 level

4. There was a statistically significant difference in means indicated by the delayed posttest between group E_2 receiving the explicit treatment and group E_1 receiving the treatment combining implicit and explicit methods in favor of the explicit treatment.

TABLE 5.--Analysis of variance of scores on the delayed posttest utilizing individual degrees of freedom.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	*F* ratio
Between	2	141.65	70.83	2.70
Exp. vs. C	1	1.37	1.37	0.05
E_1 vs. E_2	1	140.28	140.28	5.34*
Within	46	1,208.02	26.26	
Total	48	1,349.67		

*significant at the 0.05 level.

There were no other statistically significant differences.

The experimenter concluded that the completely explicit treatment E_2 was the preferred treatment for the population, treatments and tests used in this study.

There is one clear implication that can be drawn from this study. To teach mathematical proof using selected rules of logic, elementary abstract group theory or proof techniques for greatest student achievement, the communication media must clearly and explicitly identify and emphasize the materials for the student. The student should not be expected to learn subject matter that is not presented in his resource or study materials. The experimenter feels that all mathematical content should be presented by explicit means to obtain maximum student comprehension.

PROOF IN THE ELEMENTARY SCHOOL - A FEASIBILITY STUDY

Irv King
University of Hawaii
Honolulu, Hawaii

The Cambridge Conference on School Mathematics (CCSM) has recommended that the study of mathematical proof begin in elementary school (1). Although many psychologists assert that the early adolescent possesses the cognitive structures necessary for formal reasoning, there is a scarcity of empirical evidence on the subject. As stated in the Cambridge Report (2), "More information through experiment is needed."

This study was conducted in response to this need. Its purpose was to develop a unit of instruction on proof for use with sixth-grade students and, in so doing, to examine the feasibility of the CCSM proposal that proof be taught in the elementary school.

The curriculum development model constructed by Romberg and DeVault (3) was followed in developing the unit. The model is comprised of four sequential phases: Analysis, Pilot Examination, Validation, and Development. This study carried the development of a unit on proof through the first two phases of the model.

Six theorems of the kind recommended in the Cambridge Report (4) were selected for the unit:

Theorem 1: If N/A and N/B, then $N/(A + B)$.

Theorem 2: If N/A and N/B, then $N/(A - B)$.

Theorem 3: If N/A and N/B and N/C, then $N/(A + B + C)$.

Theorem 4: If $N \not/ A$ and N/B, then $N \not/ (A+B)$.

Theorem 5: If $N \not/ A$ and N/B, then $N \not/ (A - B)$.

Theorem 6: There is an infinite number of primes.

The main behavioral objectives of the unit were for the students to write valid proofs for each of these six theorems. Each behavioral objective was task analyzed. The task analytic procedure was developed by Gagne (5) to train human beings to perform complex tasks. The basic idea is to break the objective into a number of subtasks. The result is a hierarchy of behaviors which lists all of the skills which need to be learned in order to perform the desired objective. In analyzing the behaviors involved in writing a proof it was found that Gagne's task analysis model was inadequate, for it did not allow for the role which strategies play in proof. Therefore, a new, three-dimensional model was constructed and used in analyzing the proofs of the theorems.

An instructional analysis was then performed in an attempt to find effective means by which to present and motivate the various components of the unit. The basic pedagogical approach was to entice the students into generalizing the statements of the theorems from repeated instances of them. Activities with a desk computer and cartoon stories were planned to aid in motivating the unit.

With a tentative unit prepared, the investigator conducted a two-week pilot study with six sixth-grade students. An essential feature of the Romberg-DeVault model is its iterative nature: if students are unable to perform in accordance with pre-established criteria (in this case 80% of the students mastering 80% of the material), the unit is subjected to a thorough re-examination. In the first formative pilot study the students encountered difficulty with the proofs of Theorems 4, 5, and 6. Hence, the unit was rewritten and a second pilot study was conducted.

Bruner's hypothesis (6) "...that any subject can be taught effectively to any child at any stage of development" served as a guiding principle. According to Bruner, knowledge, like the cognitive structures of the mind, can be represented in many forms. If students are unable to understand a given set of ideas, it is the job of the curriculum developer to restructure those ideas into a form which is compatible with the cognitive structures of the minds of the students. In applying this principle, Theorem 6 underwent three major revisions before it was easily comprehended by the students.

After two formative pilot studies were completed, plans were made for testing the unit. Eleven detailed lesson plans with practice exercises, cartoon stories, and mastery tests were prepared. A certified elementary school teacher, whose academic preparation included six semester hours of mathematics beyond the calculus, was chosen to teach the unit. Her duties as an employee of a Research and Development Center included the teaching of experimental units.

A Nonequivalent Control Group Design (7) was used in testing the unit. An Experimental Group of ten students (mean IQ score 117) was selected from an intact classroom, and a Control Group (mean IQ score 121) was selected by matching procedures. A twenty-five item test containing both prerequisite behaviors and the proofs of the six theorems was administered to both groups prior to instruction. With the investigator serving as an observer, the teacher presented the unit to the Experimental Group. After seventeen days of instruction, the twenty-five item test was readministered to both groups. The results are summarized in Table 1.

7

TABLE 1
Pretest-Posttest Results

	Pretest		Posttest	
	Prerequisites	Proofs	Prerequisites	Proofs
Experimental	29%	0%	96%	97%
Control	30%	0%	32%	0%

An Analysis of Variance was performed on the posttest grand means and the results, as expected, were highly significant ($p < .0001$). When pretest scores, IQ scores, and/or STEP achievement scores are used as covariates, the F-ratio is increased.

The main conclusions are as follows:

1. Using the iterative curriculum development model of Romberg and DeVault, an effective unit on proof was developed. This suggests that the model is an appropriate one for developing a mathematics curriculum.

2. Under rather ideal conditions, sixth-grade students with better-than-average abilities can learn and understand the kinds of proofs recommended in the Cambridge Report.

3. The final conclusion is an unexpected one. The teacher encountered considerable difficulty in coping with the spontaneous reactions of the students. False statements went uncorrected, insightful comments were ignored, and pregnant questions were fumbled. In light of the facts that the teacher had a comparatively strong background in mathematics, that the investigator taught the unit to the teacher, and that the content of this unit represents but a small part of the total curriculum proposed in the Cambridge Report, the findings of this study indicate that the recommendations of the CCSM cannot be realized without a dramatic improvement in the mathematics competency of those reaching mathematics at the elementary school level.

REFERENCES

1. Cambridge Conference on School Mathematics. Goals for School Mathematics. New York: Houghton Mifflin Company, 1963, p. 86.

2. Cambridge Conference on School Mathematics. Goals for School Mathematics. New York: Houghton Mifflin Company, 1963, p. 15.

3. Romberg, Thomas A. and Vere DeVault. "Mathematics Curriculum: Needed Research." Journal of Research and Development in Education, Vol. 1, Number 1, Fall, 1967. Athens: The University of Georgia, 1967, p. 95.

4. Cambridge Conference on School Mathematics. Goals for School Mathematics. New York: Houghton Mifflin Company, 1963, p. 87.

5. Gagne, Robert M. and Noel E. Paradise. "Abilities and Learning Sets in Knowledge Acquisition." Psychological Monograph, 1961, Vol. 75 (14, Whole no. 518).

6. Bruner, Jerome S. The Process of Education. New York: Vintage Books, 1960, p. 33.

7. Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1968.

THE ROLE OF COUNTEREXAMPLES IN THE DEVELOPMENT OF MATHEMATICAL CONCEPTS OF EIGHTH GRADE MATHEMATICS STUDENTS

Richard James Shumway
The Ohio State University
Columbus, Ohio

This study was designed to determine whether an extensive treatment of counterexamples in the development of certain mathematical concepts in the eighth grade mathematics classroom would result in significant differences in mean scores on tests of the following factors: (a) general mathematics achievement; (b) specific mathematics achievement; (c) inductive reasoning; (d) syllogistic reasoning; (e) perceptual speed; (f) reading mathematical definitions; and (g) tendency to overgeneralize.

This study was conducted for 65 class periods from November, 1968 to March, 1969 at Marshall-University High School. The subjects for the study were 84 eighth grade mathematics students of average ability (mean Nonverbal IQ of 106) randomly assigned to four mathematics classes. Two instructors, A and B, each taught two classes. One class taught by each instructor was randomly designated the experimental class and the other the control class.

The mathematical content taught during the experiment included quadrilaterals, exponents, and operations. The experimental treatment contained an equal number of positive examples and counterexamples. The control treatment contained only positive examples, no counterexamples.

In order to determine the comparability of the classes the following pretest measures were used: The Lorge-Thorndike Verbal and Nonverbal Intelligence Tests, The Sequential Tests of Educational Progress, Mathematics, Form 3A (STEP 3A), and a unit mathematics test. An analysis of variance provided no evidence that the classes taught by Teacher A differed on the pretest measures of intelligence or mathematics achievement. An analysis of variance showed the classes of Teacher B differed significantly ($P < .05$) on the pretest measures of nonverbal intelligence and mathematics achievement.

The following posttests were administered during the seven day testing period at the end of the experiment: Inference Test (Rs-3), a mathematics achievement test (Ach-1), Figure Classification Test (I-3), Number Comparison Test (P-2), STEP 3B, Identical Pictures Test (P-3), Definitions: Operations (D-1), Generalizations: Operations (G-1), Letter Sets Test (I-1),

Nonsense Syllogisms Test (Rs-1), and Operations: Properties (Pr-1). Tests Ach-1, D-1, G-1, and Pr-1 were designed specifically for this study. Ach-1 was an achievement test, D-1 was a test of ability to read definitions of operations, G-1 was a test of the tendency to overgeneralize the properties of operations, and Pr-1 was a test of the tendency to overgeneralize the properties of operations to the basic operations of arithmetic.

On the basis of the test for the equality of the dispersion matrices, a test of the assumptions underlying the analysis, a one-way analysis of variance was chosen for the analysis for the classes of Teacher A and a two-way analysis of covariance was chosen for the analysis for the classes of Teacher B.

The analysis of the posttest results for the classes of Teacher A indicated that the mean of the experimental class was significantly higher than that of the control on Test G-1 at the five percent level. No other differences were statistically significant.

The analysis of the posttest results for the classes of Teacher B indicated that the mean of the experimental class was significantly higher than that of the control on test Pr-1 and the mean of the control class was significantly higher than that of the experimental on test P-3, both at the five percent level. No other differences were statistically significant.

For the tests for which significant differences in mean scores were found, only the differences on tests G-1 and Pr-1 were consistent over the classes of both teachers. This indicates that the use of counterexamples in teaching mathematics has a significant effect on the students' tendency to overgeneralize the properties of operations. It appears that if one would like to discourage the overgeneralization of the properties of operations the use of counterexamples is one appropriate strategy.

ALTERNATE PAPERS

To be presented in the event one of the listed papers cannot.

Speakers: 1. J. Fred Weaver, The University of Wisconsin-Madison, Madison, Wisconsin and Carole Z. Greenes, Boston University, Boston, Massachusetts, "The Ability of Pupils in Grades 4, 5, and 6 to Classify Representations of Plane Geometric Figures on the Basis of Specified Categories."

2. William P. Palow, Miami-Dade Junior College, Miami, Florida, "A Study of the Ability of Public School Students to Visualize Particular Perspectives of Selected Solid Figures."

THE ABILITY OF PUPILS IN GRADES 4, 5, AND 6 TO CLASSIFY REPRESENTATIONS OF PLANE GEOMETRIC FIGURES ON THE BASIS OF SPECIFIED CATEGORIES

J. Fred Weaver
The University of Wisconsin-Madison
Madison, Wisconsin

and

Carole Z. Greenes
Boston University
Boston, Massachusetts

It was the intent of this investigation

(1) to develop an instrument (with diagnostic features) to assess students' ability to classify representations of plane geometric figures according to specified categories, and

(2) to collect and summarize normative data from using the instrument with 4th-, 5th-, and 6th grade pupils.

The project represents refinements and extensions of explorations previously reported in the literature by the principal investigator (1, 2, 3).

An Inventory G was developed which included 25 representations of plane geometric figures and 10 classification categories. Some of the 25 representations were selected to illustrate different varieties or spatial orientations of the same class of figure: e.g.,  and  are both representations of triangles.

The Inventory G consisted of 100 items in which

(a) each of the 25 representations was to be interpreted in terms of four of the 10 classification categories, and

(b) each of the 10 classification categories was to be applied to 10 of the 25 representations. The categories were not mutually exclusive.

The set of 100 items comprising Inventory G was organized in two ways, leading to Form 2A and Form 2B which were identical in content but different in the nature of the item "stems." In Form 2A each "stem" was a representation of a plane geometric figure; in Form 2B each "stem" was a classification category. The essence of this distinction may be expressed in this way:

Form 2A	
Is this a picture of	a figure of this class?

Form 2B	
Is this class of figure	illustrated by this picture?

For either Form 2A or Form 2B there were the same three response options for each item: "Yes," or "No," and "??" (Not Sure).

A Diagnostic Record Chart was designed to record, summarize, and facilitate interpretation of a pupil's performance on Inventory G.

Performance data on Inventory G are available for 2,066 pupils of 72 teachers in grades 4, 5, and 6 of 14 elementary schools from four suburban Boston school districts,--tested about midway thru the school year.

Reliability of Inventory G

Both Form 2A and Form 2B of Inventory G were administered to pupils from one particular elementary school.

The inter-Form Pearson product-moment correlation coefficients, based upon pupils' correct responses, are summarized in the Table at the right.

	Form 2A first; then Form 2B		Form 2B first; then Form 2B	
Grade	N	r	N	r
4	26	.89	26	.82
5	19	.78	23	.79
6	22	.95	20	.85
All	67	.83	69	.82

Reports of Preliminary Explorations

1. J. Fred Weaver, "Levels of Geometric Understanding: An Exploratory Investigation of Limited Scope." Arithmetic Teacher 13: 322-32; April 1966.

2. J. Fred Weaver, "Levels of Geometric Understanding Among Pupils in Grades 4, 5, and 6." Arithmetic Teacher 13: 686-90; December 1966.

3. J. F. Weaver, "Nonmetric Geometry and the Mathematical Preparation of Elementary School Teachers." American Mathematical Monthly 73: 115-21; December 1966.

A STUDY OF THE ABILITY OF PUBLIC SCHOOL STUDENTS TO VISUALIZE PARTICULAR PERSPECTIVES OF SELECTED SOLID FIGURES

William P. Palow
Miami-Dade Junior College
North Miami, Florida

Much has been written about the ability of children to learn mathematical concepts. Bruner has stated that any subject can be taught effectively in some intellectually honest form to any child at any stage of development. Piaget and his co-workers have attempted to formulate a new theoretical framework by approaching child psychology from an epistemological viewpoint. From his many explorations Piaget has postulated a sequence for the cognitive abilities of the child. One of the vehicles of this theory is representational space or the ability to formulate images. Representational ability is seen as developing psychologically through the mathematical phases of topological, projective and Euclidean spaces in that order.

Piaget has developed and conducted many experiments to demonstrate the stages he has formulated. One of these experiments pertained to the ability to imagine cross-sections of various solids before cutting the solid. Success with this task indicated the beginnings of Euclidean spatial ability which is supposed to begin being manifested at about nine or ten years of age. Euclidean spatial abilities should, according to the model, be fully developed by age twelve.

The author extended the cross-sectioning experiment to a projective spatial task and developed an instrument to test the ability. The instrument consisted of photographs of solid figures with an arrow indicating the perspective from which the subject was to view each solid. The child was asked to imagine what the solid would look like from the direction of the arrow and to choose a response from four smaller photographs or a "none of these" choice.

One purpose of this study was to determine at what age or grade public school children attain the ability measured by this instrument. Another purpose was to take into account some of the statistical weaknesses which are so often mentioned in the literature concerning Piagetian experimentation, such as, sample size, etc. Also, the study was designed to answer questions about the effect of sex, ability level represented by I.Q. and socio-economic class on the performance on this instrument.

Two pilot studies were conducted at P.K. Younge Laboratory School, University of Florida, to determine which grades were to be included in the study and how to administer the instrument. Then, approximately one hundred twenty children were tested in each of the grades three through twelve in Duval County, Florida. The children were from two elementary schools, one junior high school and one high school. These schools were judged to have broad socio-economic backgrounds.

Data on the age, I.Q., socio-economic class and sex of the subjects were drawn from the cumulative folders of the respective students. Of the original twelve hundred students tested, 1,067 were retained for the study. The remaining were rejected because of lack of information in the cumulative folders as to I.Q. level or socio-economic background. The Kuhlmann-Anderson I.Q. test is used by Duval County to measure intelligence. This quotient and the father's occupation were recorded on most records. The Otis Dudley Duncan scale was used to rank students' socio-economic level using the father's occupation as the criterion.

The children were examined in classroom groups of about thirty students each. Before the test began, each group was shown all of the solids which individually, or in combination, composed the test items. All subjects were given ample opportunity to consider and respond to each of the twenty items on the paper and pencil examination.

The instrument was subjected to analysis for item difficulty and discrimination. Two items were rejected because the difficulty level was too low and five items were rejected because they did not discriminate the upper from the lower group. The items were then rescored.

The data were examined with analysis of covariance on the dependent variable of grade with independent variables age, sex, socio-economic level and I.Q. Both adjusted and unadjusted scores were subjected to examination by the Kramer Multiple Range Test. Analysis of covariance was also used on the dependent variable age with independent variables grade, sex, socio-economic level and I.Q. Again, both adjusted and unadjusted scores were subjected to examination by the Kramer Multiple Range Test. The data were further subjected to two three-way analyses of variance, firstly, on the variables of grade, sex and I.Q., and, secondly, on the variables of age, sex and I.Q. The socio-economic variable was deleted from the above three-way analysis of variance because earlier analyses using both analysis of variance and analysis of covariance indicated it was not significant.

For this particular sample of the population of public school children in Duval County, Florida, Piaget seems to have made a correct postulation to the effect that the ability to score well on the instrument increases with chronological age. The significance of the analyses, both on grade and age, seemed to varify the above statement. Furthermore, the results of this study indicate that the Duval children do acquire the ability at about the predicted age of twelve years.

Analysis of the data indicates that boys scored better than girls on this instrument in this sample. Although the boys seem to have an advantage at this task, at the 5 percent level, the difference is not significant at the 1 percent level.

In this study, ability level as determined by an I.Q. score on the Kuhlmann-Anderson Test made a great difference on the ability to perform well on this test. The higher I.Q. group had a great advantage over the lower I.Q. group.

For this particular sample, each of the three socio-economic classes performed just as well on the instrument. Very little difference could be found between groups and no difference was found that could not be due to chance.

It would seem that for this sample that the high mental ability, older boys in the higher grades in the public schools in Duval County were the best achievers on the instrument of this study.